国际电气工程先进技术译丛

WILEY

固体氧化物燃料电池的动态建模与预测控制

Dynamic Modelling and Predictive Control in Solid Oxide Fuel Cells: First Principle and Data-based Approaches

[加拿大] Biao Huang
Yutong Qi
AKM Monjur Murshed 著

孙玉绣 汪浩 等译



机械工业出版社
CHINA MACHINE PRESS

VISIT…

LANZAROTE Caliente.com

关于本书

在目前的能源市场中，高温固体氧化物燃料电池(SOFC)被认为是主要的燃料电池技术的竞争者之一。然而,为了操作作为一个高效的发电系统,SOFC需要一个适当的控制系统，还需要一个详细的过程动态模型。为了能够论述最先进的动态建模、估计和SOFC系统的控制，本书介绍了原始建模方法以及由作者研究的全新的成果。本书通过采用基于动态建模和基于数据的方法,并考虑控制的各个方面,包括建模、系统辨识、状态估计、传统和先进的控制，全面覆盖和SOFC技术的许多方面。

国际电气工程先进技术译丛

固体氧化物燃料电池的动态建模与预测控制

Biao Huang
〔加拿大〕 Yutong Qi 著
AKM Monjur Murshed
孙玉绣 汪 浩 等译

机 械 工 业 出 版 社

在目前的能源市场中，高温固体氧化物燃料电池（SOFC）被认为是主要的燃料电池技术的竞争者之一。然而，为了操控一个高效的发电系统，SOFC 需要一个适当的控制系统，还需要一个详细的建模过程动态。为了能够论述最先进的动态建模、估计和 SOFC 系统的控制，本书介绍了原始建模方法以及由作者开发的全新的成果。本书通过采用基于动态建模和基于数据的方法，并考虑控制的各个方面，包括建模、系统辨识、状态估计、传统和先进的控制，全面覆盖和 SOFC 技术的许多方面。

本书介绍的方法适合学习化学工程的基本原理、系统辨识、状态估计和过程控制，因此，它适用于化工、机械、电力、电气工程，特别是在过程控制、过程系统工程、控制系统、燃料电池方面的研究生与教师以及研究和工程技术人员参阅。本书还能帮助研究人员了解相关的基础知识以及当前在 SOFC 动态建模和控制技术方面的概述。

译 者 序

为了缓解当今日益加重的能源危机与环境压力，开发和利用新能源和可再生能源是当今国际一大研究热点。燃料电池作为一种清洁能源受到了人们的极大关注。燃料电池是一种直接将化学能转化为电能的电化学装置。在各种燃料电池中，低温质子交换膜燃料电池（PEMFC）和高温固体氧化物燃料电池（SOFC）已被确认为有前途的燃料电池技术，并将在未来市场中占有很大份额。为了达到操作和控制SOFC 系统的目的，对 SOFC 动态特性的建模和仿真研究是十分必要的。

建模和控制器设计是两个必要的先进过程控制策略的组成部分，两者错综复杂、相互依存。从过程控制的角度来看，模型应该可以用来设计控制器，但也要足够详细地提供系统动态真实模型解释。

动态模型可以用来研究在不同操作条件下燃料电池的响应，以便说明与设计和材料选择有关的缺陷。通过最优化控制，可以引导操作条件朝向有利于提高燃料电池耐用性和效率的方向进行。因此，动态建模和控制在燃料电池的发展中是十分必要的。随着廉价计算机的出现，工业上也采用复杂控制器，如非线性模型预测控制器。

在本书中，诸位作者参考了大量的期刊论文、会议论文集资料以及作者研发的全新科研成果，概述了最先进的动态建模、估计和 SOFC 系统的控制的最新发展。

我国是一个能源消费大国，能源紧张问题尤为突出，更迫切需要开发高效利用燃料的相关技术和方法。尽管我国在 SOFC 方面已经取得了一些成绩，但是与发达国家相比，该领域相关研究起步晚、资金投入不够、技术积累不多，所以在此领域技术明显落后。在此背景下，翻译本书的目的是向国内读者系统介绍基于动态建模和基于数据的方法，考虑控制的各个方面，包括建模、系统辨识、状态估计、传统和先进的控制，以及当前在 SOFC 动态建模和控制技术方面的概述，为研究人员在过程系统工程与控制系统工程方面探索新的控制理论和算法提供良好的参考。

我们期望本书的出版对国内相关领域的发展有所帮助。如果能做到这一点，我们将非常欣慰和鼓舞。

本书的翻译人员有：孙玉绣（第 1 ~ 6 章）、汪浩（第 7 ~ 12 章）。田晓萌、张欣、赵博文、沈强、付丛丛、邓思旭、吴春卉、王先明、曹以靓、靳晴、郭超、郭振玺、王鲁、刘炳毅、古悦、邵可然、邬浩等人参与了本书部分翻译工作，全书由孙玉绣统稿。机械工业出版社顾谦老师为本书的翻译、编辑、加工、出版付出了辛勤的劳动，在此深表谢意。由于时间仓促以及我们的知识面有限，译文中肯定有疏漏或欠妥之处，诚冀读者批评指正。

译 者

原 书 前 言

为了缓解当今日益加重的能源危机与环境压力，开发和利用新能源和可再生能源是当今国际一大研究热点。严禁排放温室气体、含硫氧化物和含氮氧化物气体相关环境法规的颁布，很大程度上缩短了寻找一种清洁能源的进程。燃料电池作为一种清洁能源可替换来源受到了人们的极大关注。燃料电池是一种直接将化学能转化为电能的电化学装置。由于其不涉及任何的转动或热组件，燃料电池没有任何摩擦和燃烧损失。此外，电池中未使用的燃料可以用来产生更多的能量，使其达到高的总效率。

在各种燃料电池中，低温质子交换膜燃料电池（PEMFC）和高温固体氧化物燃料电池（SOFC）已被确认为有前途的燃料电池技术，并将在未来市场中占有很大份额。

为了达到操作和控制 SOFC 系统的目的，对 SOFC 动态特性的建模和仿真研究是十分必要的。建模和控制器设计是两个必要的先进过程控制策略的组成部分，两者错综复杂、相互依存。从过程控制的角度来看，模型应该可以用来设计控制器，但也要足够详细地提供系统动态真实模型解释。

本书着重介绍两种模型用于 SOFC 研究，分别为第一性原理和数据基模型。第一性原理模型考虑电化学、化学和热力学方面，并提供一组非线性常微分方程（ODE）。用于燃料电池系统组件（如换热器、重整装置和燃烧器）的零维热模型也提供了燃料电池系统的仿真和控制。同时，数据基模型是通过系统辨识方法研发的。

动态模型用来研究在不同操作条件下燃料电池的响应，以便说明与设计和材料选择有关的缺陷。通过最优化控制，可以引导操作条件朝向对有利于提高燃料电池耐用性和效率的方向进行。因此，动态建模和控制在燃料电池的发展中是十分必要的。

随着廉价计算机的出现，工业上也采用复杂控制器如非线性模型预测控制器。在本书中，常规控制器和先进的模型预测控制器均应用在燃料电池系统中。

本书试图将作者在过去八年里发展或出版的成果完整地整合在一本书中，就这一点来说，本书将适合作为研究生及科研人员的教科书或专题著作，为寻找与动态模型和控制及应用相关的基础知识以及先进的技术提供帮助。本书介绍了几个独立的燃料电池模型，并附有详细的参数和解释，为研究人员在过程系统工程与控制系统工程方面寻找具有挑战性的问题来测试新的控制理论和算法提供良好的参考。本书适合从事化学工程、机械工程、电气工程等相关的过程控制、燃料电池、过程系统工程或控制系统等专业的研究生和科研人员阅读，还可以用作教材或者作为参考，有助于学习化学工程的基本原理、系统辨识、状态估计和过程控制。另外，本书附有燃料电池应用实例的生动图示。

原 书 致 谢

我们要特别感谢我们的同行和合作者，Kumar Nandakumar 教授、Karl Chuang 以及 Jingli Luo，他们与我们进行了很多次讨论，让我们了解到在过去的几年里燃料电池的主要研究成果，同时，我们也感谢加拿大阿尔伯塔大学计算机程序控制团队，他们为过程控制研究提供了模拟环境。阿尔伯塔大学包括化学和材料工程专业的广泛智囊团为本书的出版提供了不同思想的交融和发展。我们非常感激我们的工业合作者，他们为我们提供了过程控制研究的相关实践的数据。我们也感谢我们的计算支撑团队以及阿尔伯塔化学和材料功能专业的其他帮助者。感谢加拿大自然科学与工程研究委员会（NSERC）和西加拿大燃料电池创制权（Western Canada Fuel Cell Initiative，WCFCI）对相关研究工作的支持。最后同样重要的是，我们要感谢 Shima Khatibisepehr 和 Seraphina Kwak 对本书的详细评阅和对很多章节的建议。

目　　录

译者序

原书前言

原书致谢

第 1 章　绪论 ······ 1

1.1　燃料电池技术概述 ······ 1

1.1.1　燃料电池种类 ······ 1

1.1.2　板式和管式设计 ······ 2

1.1.3　燃料电池体系 ······ 3

1.1.4　燃料电池的优、缺点 ······ 4

1.2　建模、状态估计和控制 ······ 4

1.3　本书覆盖范围 ······ 5

1.4　本书大纲 ······ 5

第 1 部分　基 本 原 理

第 2 章　化学反应的第一性原理建模 ······ 7

2.1　热力学 ······ 7

2.1.1　能量的形态 ······ 7

2.1.2　第一定律 ······ 7

2.1.3　第二定律 ······ 8

2.2　热传递 ······ 9

2.2.1　传导 ······ 9

2.2.2　对流 ······ 10

2.2.3　辐射 ······ 11

2.3　质量传递 ······ 13

2.4　流体力学 ······ 15

2.4.1　黏性流 ······ 15

2.4.2　速度分布 ······ 16

2.4.3　伯努利方程 ······ 16

2.5 变化方程式 …… 16
2.5.1 连续性方程 …… 17
2.5.2 运动方程 …… 17
2.5.3 能量平衡方程 …… 18
2.5.4 连续性方程的种类 …… 19
2.6 化学反应 …… 20
2.6.1 反应速率 …… 20
2.6.2 可逆反应 …… 21
2.6.3 反应热 …… 21
2.7 注解和参考文献 …… 22

第 3 章 系统辨识 I …… 23

3.1 离散时间系统 …… 23
3.2 信号 …… 28
3.2.1 输入信号 …… 28
3.2.2 信号的光谱特性 …… 32
3.2.3 输入信号的持续激励 …… 35
3.2.4 输入设计 …… 39
3.3 模型 …… 39
3.3.1 线性模型 …… 39
3.3.2 非线性模型 …… 43
3.4 注解和参考文献 …… 45

第 4 章 系统辨识 Ⅱ …… 46

4.1 回归分析 …… 46
4.1.1 利用外生输入模型的移动平均自动回归 …… 46
4.1.2 线性回归 …… 47
4.1.3 线性回归分析 …… 48
4.1.4 加权最小二乘法 …… 49
4.2 预测误差方法 …… 52
4.2.1 最优预测 …… 53
4.2.2 预测误差方法 …… 57
4.2.3 独立参数预测误差法 …… 60
4.2.4 PEM 渐近方差属性 …… 61
4.2.5 非线性辨识 …… 62
4.3 模型验证 …… 64

4.3.1　模型结构选择 …… 64
4.3.2　简约原则 …… 65
4.3.3　模型结构比较 …… 66
4.4　经验法 …… 67
4.4.1　非零处理 …… 67
4.4.2　漂移干扰处理 …… 68
4.4.3　鲁棒性 …… 68
4.4.4　额外的模型验证 …… 68
4.5　闭环辨识 …… 69
4.5.1　直接闭环辨识 …… 70
4.5.2　间接闭环辨识 …… 71
4.6　子空间辨识 …… 75
4.6.1　符号 …… 76
4.6.2　回归分析法子空间辨识 …… 80
4.6.3　范例 …… 82
4.7　注解和参考文献 …… 85

第5章　状态估计 …… 86

5.1　随机动态系统过滤技术进展 …… 86
5.2　问题界定 …… 88
5.3　状态估计的序贯贝叶斯推理 …… 89
5.3.1　KF 和 EKF …… 92
5.3.2　UKF …… 94
5.4　范例 …… 97
5.5　注解和参考文献 …… 100

第6章　模型预测控制 …… 102

6.1　最先进的模型预测控制 …… 102
6.2　基本原理 …… 103
6.2.1　MPC 模型 …… 104
6.2.2　自由和强迫响应 …… 105
6.2.3　目标函数 …… 105
6.2.4　限制条件 …… 106
6.2.5　MPC 规则 …… 106
6.3　DMC …… 107
6.3.1　预测 …… 107

6.3.2　无变形控制移动的 DMC …… 109
6.3.3　变形控制移动的 DMC …… 109
6.3.4　DMC 算法反馈 …… 110
6.4　NMPC …… 113
6.5　NMPC 通用优化准则 …… 115
6.6　离散模型：正交配置法 …… 115
6.6.1　预测时域 1 的正交配置方法 …… 116
6.6.2　预测时域 *N* 的正交配置方法 …… 118
6.7　MPC 的优、缺点 …… 120
6.8　最优化 …… 121
6.9　范例：混沌系统 …… 122
6.10　注解和参考文献 …… 123

第 2 部分　管式 SOFC

第 7 章　管式 SOFC 动态模型：第一性原理方法 …… 124

7.1　SOFC 堆栈的设计 …… 124
7.2　转换过程 …… 125
7.2.1　电化学反应 …… 125
7.2.2　电流动态 …… 128
7.3　扩散动态 …… 129
7.3.1　扩散传递函数 …… 130
7.3.2　简化的扩散传递函数 …… 131
7.3.3　扩散动态模型 …… 132
7.3.4　扩散系数 …… 133
7.4　燃料输送过程 …… 134
7.4.1　重整/转移反应 …… 134
7.4.2　传质过程 …… 135
7.4.3　动量传递 …… 137
7.4.4　能量转移和热交换 …… 138
7.5　空气输送过程 …… 139
7.5.1　阴极通道传质 …… 139
7.5.2　阴极通道动量传递 …… 140
7.5.3　阴极通道能量传递 …… 140
7.5.4　注入通道空气 …… 140

7.6 SOFC 温度 …… 141
7.6.1 动态能量交换过程 …… 141
7.6.2 热传导 …… 142
7.6.3 对流 …… 143
7.6.4 辐射 …… 144
7.6.5 电池温度模型 …… 145
7.6.6 注入管温度模型 …… 145
7.7 最终动态模型 …… 146
7.7.1 I/O 变量 …… 146
7.7.2 状态空间模型 …… 147
7.7.3 模型验证 …… 150
7.8 模拟动态属性研究 …… 152
7.8.1 扩散动态 …… 152
7.8.2 燃料输送过程动态 …… 154
7.8.3 空气输送过程动态 …… 156
7.8.4 外部负载动态 …… 157
7.9 注解和参考文献 …… 160

第 8 章 管式 SOFC 动态模型：简化的第一性原理方法 …… 161

8.1 简介 …… 161
8.1.1 过程变量的关系 …… 161
8.1.2 功率输出限制 …… 162
8.2 SOFC 堆栈的低阶状态空间模型 …… 162
8.2.1 物理过程 …… 163
8.2.2 建模假设 …… 163
8.2.3 I/O 变量 …… 163
8.2.4 电压 …… 165
8.2.5 分压 …… 166
8.2.6 流速 …… 167
8.2.7 温度 …… 169
8.3 非线性状态空间模型 …… 170
8.4 仿真 …… 172
8.4.1 验证 …… 172
8.4.2 输入阶跃响应 …… 174
8.4.3 干扰阶跃响应 …… 174
8.5 注解和参考文献 …… 177

第 9 章　管式 SOFC 的动态建模与控制：系统辨识方法 …… 178

9.1　简介 …… 178
9.2　系统辨识 …… 178
9.2.1　变量选择 …… 178
9.2.2　阶跃响应测试 …… 179
9.2.3　非典型性阶跃响应 …… 181
9.2.4　输入设计 …… 183
9.2.5　线性系统辨识 …… 184
9.2.6　非线性系统辨识 …… 199
9.3　PID 控制 …… 203
9.3.1　设定点跟踪 …… 204
9.3.2　抗干扰 …… 204
9.3.3　离散时间过程 IMC …… 206
9.3.4　多环控制 SOFC 离散时间 IMC 应用 …… 214
9.4　闭环辨识 …… 218
9.5　注解和参考文献 …… 223

第 3 部分　板式 SOFC

第 10 章　板式 SOFC 动态建模：第一性原理方法 …… 224

10.1　简介 …… 224
10.2　几何学 …… 225
10.3　堆栈电压 …… 225
10.4　质量守恒 …… 227
10.5　能量守恒 …… 228
10.5.1　集中模型 …… 228
10.5.2　详细模型 …… 229
10.6　仿真 …… 232
10.6.1　稳态响应 …… 233
10.6.2　动态响应 …… 234
10.7　注解和参考文献 …… 236

第 11 章　板式 SOFC 系统动态建模 …… 237

11.1　简介 …… 237

11.2 燃料电池系统 …… 237
11.2.1 燃料和空气热交换 …… 238
11.2.2 重整装置 …… 239
11.2.3 燃烧器 …… 240
11.3 连同电容器的 SOFC …… 241
11.4 仿真结果 …… 242
11.4.1 燃料电池系统仿真 …… 242
11.4.2 超级电容器 SOFC 堆栈 …… 245
11.5 注解和参考文献 …… 245

第 12 章 板式 SOFC 系统的模型预测控制 …… 246

12.1 简介 …… 246
12.2 控制目标 …… 247
12.3 状态估计：UKF …… 248
12.4 稳态经济优化 …… 250
12.5 控制和仿真 …… 251
12.5.1 LMPC …… 252
12.5.2 NMPC …… 254
12.5.3 优化 …… 254
12.6 结果与讨论 …… 256
12.7 注解和参考文献 …… 257

附录 性质和参数 …… 258

A.1 参数 …… 258
A.2 气体性质 …… 259

参考文献 …… 263

第1章　绪　　论

燃料电池是直接将化学能转化为电能的电化学装置。由于燃料电池中能量的产生不涉及任何移动部件及其不同于热引擎的工作原理，所以由这些燃料电池产生的能量既不产生任何机械摩擦损失也不受卡诺循环效率的限制。而且，电池中未反应的燃料可以用来产生更多的电力。增加电池的总效率还可以通过从废气中重新产生热量的方式实现。

解决当前能源危机的效率路径是寻找一个替代能源以便代替现有的不断消耗的常规能源。严格的环境法规已限制 SO_x 和 NO_x 等温室气体的排放，因此缩小了寻找清洁能源的范围。以上是燃料电池作为一种清洁能源的替代来源，其开发利用正日益受到人们关注的主要原因。

然而，燃料电池作为主要能源在商业化应用方面存在许多障碍，主要的障碍来自于高的生产成本。大量的研究集中在设计和操作燃料电池以便降低生产成本，将这些设备变成可行且有竞争力的能源上。电解质、催化剂以及电极材料的选择也会影响到燃料电池的成本。许多研究人员致力于研究这个领域。在不同操作条件下模拟燃料电池系统以便弥补与设计和材料选择相关的所有缺陷是十分有必要的。从不同的角度来看，模型和模拟范围可以达到从微观到系统的水平。本书着重从安全、高效运行的燃料电池系统的过程控制分析固体燃料电池作为一种动力来源，主要包括相关模型控制、状态评估和控制器设计。

1.1　燃料电池技术概述

一个燃料电池结构单元是由三部分组成的：电解液、阴极和阳极。燃料不断地被送入燃料电池的阳极，合适的氧化剂，通常是空气，被送入到阴极。电解液的主要作用是防止与阳极和阴极同时相连接的氧化剂与燃料直接接触。电解液也能够保证氧化剂或还原剂离子到达另一边参与电化学反应。

1.1.1　燃料电池种类

基于电解质和燃料的选择不同，将燃料电池进行分类如下：

● 固体氧化物燃料电池（SOFC）：SOFC 使用固体陶瓷型氧化物，因此这样命名。Y_2O_3 稳定型 ZrO_2（YSZ）是一种常见的使用在 SOFC 中的电解质，操作温度通常较高（600～1000℃）。由于电解质和电极都是固态的，所以 SOFC 的设计和制造采用最通用的方式，包括板式和管式设计。

● 熔融碳酸盐燃料电池（MCFC）：MCFC 采用不同种类碳酸盐相结合作为电

解质，这些碳酸盐通常是陶瓷基材料，其操作温度也比较高，通常在600～700℃。

● 质子交换膜燃料电池（PEMFC）：在此类电池中，高分子离子交换膜被用作电解质，操作温度通常较低，40～80℃。

● 磷酸燃料电池（PAFC）：在PAFC中的电解质为100%磷酸，其保存在碳化硅结构中。操作温度大约为150～220℃，该操作温度使这种燃料电池具有一定的优势，操作温度有利于灵活设计燃料电池和设备平衡（BOP）。

其他种类的燃料电池包括：碱性燃料电池（AFC）、直接甲醇燃料电池（DMFC）、再生燃料电池（RFC）、金属空气燃料电池（MAFC）。燃料（2004）为不同类型的燃料电池存在的主要差异进行了总结。操作温度低的PEMFC和温度高的SOFC被认定为有前景的燃料电池技术，它们将成为未来市场的重要角色。图1.1所示为典型的氢固体氧化物燃料电池的基本原理，电池内部进行的化学反应，直接参与电力的产生，如下所示：

$$\begin{aligned}&\text{在阳极三相界面（tpb）：} H_2 + O^{2-} \rightarrow H_2O + 2e^- \\ &\text{在阴极三相界面（tpb）：} 1/2\ O_2 + 2e^- \rightarrow O^{2-}\end{aligned} \tag{1.1}$$

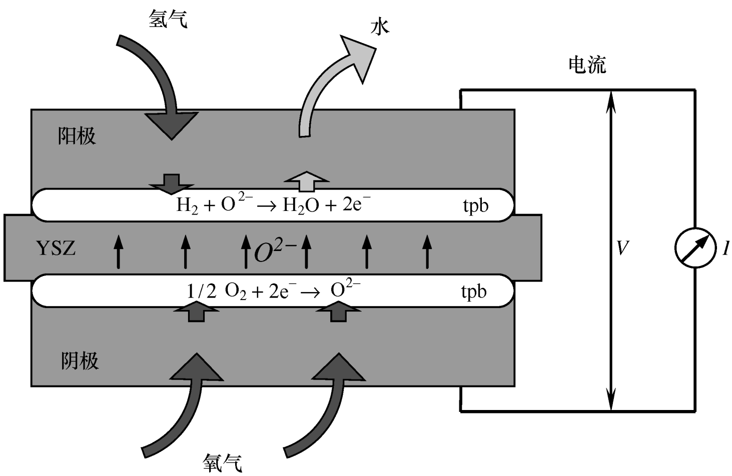


图1.1　固体燃料电池原理

在SOFC的阳极，氢气与通过电解质迁移的氧离子反应，生成水并释放电子。在阴极，氧气和电子反应生成氧离子，O^{2-}通过电解质运送到阳极。电子在阳极产生，流经外部电路，然后到达阴极。因此，这些反应持续进行流经外部电路，供应电力。通常燃料电池工作温度较高，在600～1000℃，以满足电解质离子电导率的要求。

氢被用作燃料电池的燃料，通常由天然气的蒸气重整产生。对于操作温度较高的燃料电池，在电池阳极内的重整反应可以在内部进行。

1.1.2　板式和管式设计

为了满足大部分应用程序的电压需求，燃料电池必须由成堆的单个电池连接。

一个 SOFC 堆栈由许多 SOFC 电池组成，以产生高的电压输出。在设计 SOFC 堆栈和电池的时候，一些因素必须考虑：如气体输送、热应力、机械力、固有电和离子阻力以及密封材料的选择。燃料电池可被制成各种几何形状，最常见的有板式和管式两种设计，如图 1.2 和图 1.3 所示。

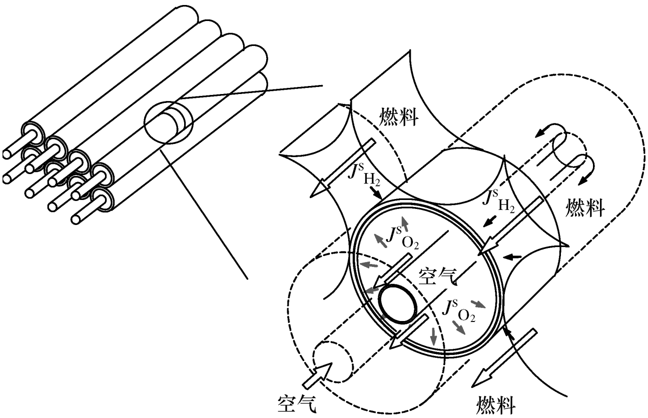


图 1.2　SOFC 堆的板式设计和电池

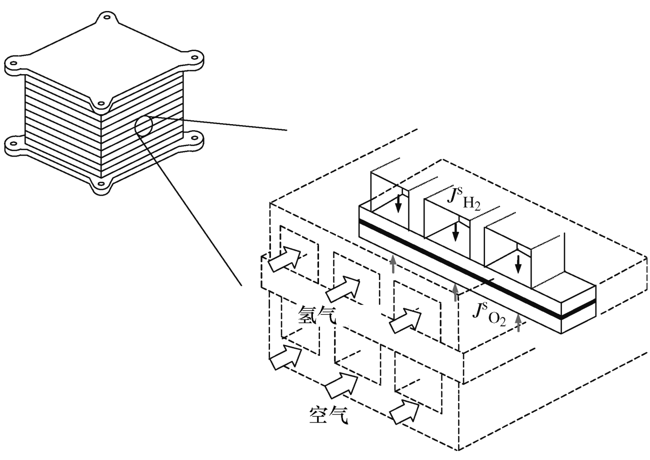


图 1.3　SOFC 堆的管式设计和电池

管式设计的主要优点在于其不需要密封来分离燃料和空气流。另一个优点在于管式的形状能够提高电池的强度。管式的形状还能提高气体输送性能。这种管式设计适合于固定的大规模发电应用领域。

另一方面，板式 SOFC 设计的主要优点在于它的电阻率低，板式 SOFC 更加适用于移动和低功耗应用领域。

1.1.3　燃料电池体系

在理想的燃料电池中，氢气被用作燃料，空气被用作氧化剂。这样工作的燃料

电池可作为最清洁的可能能源，其副产物为水。然而，氢气在自然界中含量不高，实际上作为燃料电池的氢气燃料需要从其他物质中产生。富氢燃料通常被用来生产氢气，无论在内部还是尾部的燃料电池。因此，燃料电池装置通常包括预处理和后处理反应物和生成物。也可被称作 BOP 的组件，可包括压缩机、涡轮机、用于重整燃料的热交换器以及 DC-AC 变换器或连接现有电网和燃料电池的逆变器。

压缩机或者鼓风机可保证建立必要的压力使得反应物和产物通过不同的组件，燃料电池中未反应的燃料可通过一个燃气轮机产生更多的电力。压缩机因此提供了一个净功和燃料电池本身产生的直接电力。在住宅建筑物应用领域，热排出的气体可以用来供应热水，为家庭提供热量。

燃料电池直接把化学能转化为电能，输出的是直流电压，适用于小设备的运转。对于燃料电池发电厂，直流电源需要变换成交流电，以满足电网需求，因此，电池平衡设备还包括功率调节装置（PCU）。

1.1.4　燃料电池的优、缺点

与传统的发电系统（如电池和发电机）相比，燃料电池具有很多优势。与其他能源一样，燃料电池也有一些优点和缺点，如下所述：

优点：

● 不像涡轮机，燃料电池系统没有任何移动的零件，因此没有任何机械摩擦及相关联的损失。操作环境安静，且维护少。

● 与热力发动机相比，燃料电池将化学能直接转化为电能。因此其不受卡诺循环效率限制。

● 燃料电池中的排气（未反应的燃料）气体可以用来产生超额电力与热机耦合，因此，增加其效率。

● 燃料电池的效率并不受尺寸的限制，因此，小型燃料电池可驱动一台笔记本电脑或个人电子产品，可产生能源的效率相当于一个 10MW 燃料电池发电厂。

● 燃料电池的燃料取材广泛。

● 反应限制在一个燃料电池内部，只有离子之间反应，其限制了 NO_x 和 SO_x 气体释放到环境中。

缺点：

● 时至今日，燃料电池相比于其他能源的生产技术，价格昂贵。

● 大多数燃料电池使用氢作为燃料，由于成本和与生产、存储、氢运输相关的复杂性，这阻碍了这些设备的商业化进程。

● 与电池相比，燃料电池功率密度较低，具有较短的寿命。

● 燃料气体的杂质可能毒害催化剂电极。

1.2　建模、状态估计和控制

建模、状态估计和控制器的设计，构成了先进的流程控制技术，它们复杂地相

互依赖。例如，建模（无论采用第一性原理还是基于数据的，线性或非线性，零维或三维模型）影响设计的控制器和状态估计技术。因此，一个过程的模型应该是基于客观的。为了达到控制的目的所设计的简单模型或许比复杂的三维模型更好，另一方面，或许适合过程的设计和性能分析。简言之，本书建模的目标是寻找适合控制设计的模型设计。

同样，控制器设计和状态估计技术应有目标性。一个复杂模型表达的过程可能足够稳定来控制一个常规比例积分微分（PID）控制器。另外，一个简单过程可能有很多环境和经济约束，需要一个多变量控制器来维持系统的最优性能。

SOFC 系统展现出高度的非线性特征，需要研究各种建模、估计和控制技术。本书涵盖了所有这三个相互关联的方面，即建模、状态估计和控制。

1.3　本书覆盖范围

本书由三部分组成。第 1 部分提供了一个基本原则的教程，该教程在后续章节中将会应用到。具体来说，化学工程原理、系统辨识、状态估计和模型预测控制均应用到燃料电池系统，因此它们的原理也将在第 1 部分中加以介绍。

第 2 部分着重详细简化介绍管式 SOFC 动态模型。第一性原理建模设计考虑所有气流动力学，包括质量、能量和力矩平衡。

第 2 部分的基于数据的建模基于系统辨识的方法，该方法已在第 1 部分详细介绍过。系统辨识的各个方面阐述了通过燃料电池造型的应用。作为辨识系统的自然结果，模型辨识用于反馈控制设计包括 PID 和 IMC。

燃料电池的最近研究进展和日益增加的关注，大大促进了对燃料电池的模型以及系统组件的研究。这些模型范围从零维模型到复杂的三维模型，还包括性能评价的区域和燃料电池的优化设计。然而，很少涉及发展系统水平相关模型的控制，只是充分描述燃料电池系统动力学，是足够简单的控制设计。这促使我们在第 3 部分内容中，开发燃料电池的集总模型和形成燃料电池系统的平衡组件。为了达到多元化的叙述，本书第 3 部分致力于介绍板式 SOFC。

广泛的线性和非线性控制技术已经成功应用在各种行业，特别是在过去的十年里，随着廉价计算力的出现，传统的 PID 控制器已向先前非可实现的控制器转变，如非线性模型预测控制器（NMPC），这促使我们去尝试 SOFC 系统的 NMPC，优化 SOFC 系统电能最大化。

1.4　本书大纲

本书编写布局如下：

第 2 章介绍了化学工程基本原理，这个原理是随后几章中模型第一性原理的

基础。

第 3 章介绍系统辨识的基础知识，包括过程的离散时间表示、信号、输入信号和数据基模型的结构。

第 4 章介绍系统辨识的先进主题，包括误差预报法、非线性辨识、模型验证、实践问题研究、闭环辨识和子空间辨识。

第 5 章介绍状态估计方法以及参数估计。本章的重点是无迹卡尔曼滤波（Unscented Kalman Filter），在随后的几章中将应用无迹卡尔曼滤波在 SOFC 的模型预测控制研究中。

第 6 章提供在两个线性和非线性模型预测控制（MPC）上的教程概述。按照现有的 MPC 在工业实践上的应用，也讨论了经济优化策略。

第 7 章介绍了详细的第一性原理方法应用在管式 SOFC 动力学模型建立上。这一章详细说明了在第 2 章讨论的关于如何将化工原理应用到解决燃料电池过程中的模型问题。

第 8 章应用化学工程原理得到用于控制设计的降阶动态模型。这一章也可以作为独立的章用来介绍 SOFC 操作原则和简化的建模过程。

第 9 章详细说明为什么系统辨识方法可以应用于解决数据基 SOFC 的建模问题，以确定的模型为基础，设计传统的反馈控制器以及模拟管式 SOFC。

第 10 章详细介绍第一性原理方法在板式 SOFC 动态模型中的建立应用。

第 11 章介绍考虑到整个燃料电池系统（包括 BOP）的第一性原理模型，以及组件如何组成一个系统。

第 12 章说明了 SOFC 系统的模型预测控制设计，包括状态估计、线性和非线性 MPC。

根据读者的背景和兴趣，读者可以根据章节的序列阅读整本书，也可以选择性地阅读一些章节。在后一种情况下，基于读者的知识背景，建议阅读材料的线路如下：

1. 对第一性原理为基础的建模感兴趣的读者，阅读序列如下：

第 2 章→第 7、8 章→第 10、11 章

2. 对第一性原理为基础的建模和控制内容感兴趣的读者，阅读序列如下：

第 2 章→第 5 ~ 8 章→第 10 ~ 12 章

3. 对数据基系统辨识方法内容感兴趣的读者，阅读序列如下：

第 3 ~ 5 章→第 9 章

对于其他补充材料，参考本书的官方首页：http：//www.ualberta.ca/~bhuang/SOFCbook/。

第 1 部分　基本原理

第 2 章　化学反应的第一性原理建模

2.1　热力学

热力学一词源自希腊语“therme”（热）和“dynamics”（动力），早期用于描述研究热转化为功。直到第一次成功解释了大气蒸汽机，热力学才成为一门科学。现在，它被广义为包括能量与能量转换的所有方面。

自然界中最基本的法则之一是能量守恒。能量既不能被创造也不能被消灭。它只能从一种形式转换成另一种或从一个地方运输到另一个地方。热力学第一定律解释了能量守恒，并且强调能量是热力学性能。热力学第二定律强调能量有数量特性，且实际过程的发生方向是能量数量减少的方向。

2.1.1　能量的形态

能量以多种形态存在，如热能、机械能、动能、势能、电能、磁能、化学能和核能。通常，各种形态的能量可以分为两类：宏观和微观。系统拥有的能量形态作为一个整体与外部参照物被归结为宏观能量，如动能和势能。另一方面，能量的形态与系统的分子结构和分子运动相关的是微观能量，如热能、化学能和核能。一个系统所有微观能量的总和称为该系统的内能，U。

一部分内能与分子动能和分子间引力结合的归类为热或热能。内能与分子中的原子键结合的称为化学（或键）能。内能与原子核的键相结合的称为核能。

不被包含在系统中的能量形态有化学能和核能，可以被认为是能量的动力形态。能量的动力形态通常在系统的边界出现，它们代表在过程中系统能量的获得或失去。系统可通过质量转移而进行质量交换，因为有时质量会转入或转出系统；能量包含在质量中并随之转移。即使没有质量转移，能量的相互作用也会通过热转移和功发生。在一个密闭系统中，热转移和功是能量相互作用的仅有的两种形式。

2.1.2　第一定律

一般，热力学第一定律解释了能量守恒定律，它提供了一个学习各种形态的能量和能量相互作用关系的基础。在 19 世纪上半叶以焦耳的实验为基础，第一定律指出能量既不能被创造也不能被消灭，它只能转换形式。在没有热转移的情况下，

一个系统从一个特定状态到另一个特定状态经历一系列的绝热过程，所做的净功与不管自然过程的封闭系统相同，并且过程的细节相似。同样的，在系统和周围环境不做功的情况下，经热传递的数量与封闭系统全部能量的改变量相等。所以密闭系统的第一定律可以表达为

$$Q - W = \Delta U + \Delta \mathrm{KE} + \Delta \mathrm{PE} \tag{2.1}$$

式中，Q 是周围环境传递给系统的净热；W 是系统所做的净功；ΔU 是内能的改变值；$\Delta \mathrm{KE}$ 和 $\Delta \mathrm{PE}$ 分别是动能和势能。

内能通过物质的强度性质表示了它的宏观性能、温度和压强。温度是物质的物理性质，定量地表达了与冷和热相同的概念。强度性质阐释了物质的组成颗粒无规则运动的活动并且是颗粒运动的结果。压力是通过气体和液体组成颗粒的运动来衡量力的存在的。

材料内能的改变与它的温度相关的材料性能称为比热，C。比热的定义是单位质量的某种物质，温度升高 1℃ 所吸收的热量。需要区分两种比热：恒容比热 $\boldsymbol{C}_{\mathrm{v}}$ 和恒压比热 $\boldsymbol{C}_{\mathrm{p}}$。气态物质的这些性质会有很大不同。

恒容比热的定义是

$$C_{\mathrm{v}} = \left(\frac{\partial U}{\partial T}\right)\bigg|_{\mathrm{v}} \tag{2.2}$$

恒压比热的定义是

$$C_{\mathrm{p}} = \left(\frac{\partial H}{\partial T}\right)\bigg|_{\mathrm{p}} \tag{2.3}$$

另一个重要性质是焓，H，定义是

$$H = U + PV \tag{2.4}$$

焓是两个性质的组合：内能 U 和 PV 的乘积，即通过媒介从环境交换能量并确定体积与压强。焓是衡量动力学系统的全部能量，并广泛应用于工程实际。

对于理想气体，温度、压强和容积与理想气体方程有关：

$$PV = nRT \tag{2.5}$$

式中，n 是摩尔数；R 是气体常数。

2.1.3 第二定律

一个常识是一杯热咖啡放在房间里就凉了，咖啡不可能自发地从冷的房间吸收热并给自己加热。这一例子说明发生过程有一个必然的方向。热力学第一定律对发生过程的方向没有限制，并且不能确定此过程是否会发生。而热力学第二定律可以确定过程是否会发生。

凯尔文-普朗克（Kelvin-Planck）对热力学第二定律的表述如下：

对任意的循环运转装置，不可能从单一热源取出热使之完全变成功，而不发生其他变化。

这意味着一个热机循环必须与低温槽交换热，同时一个高温源使其能连续

运转。

Clausius 关于热力学第二定律的表述如下：

对任意的循环运转装置，不可能使热从低温物体传给高温物体，而不引起其他变化。

可以等同理解为热没有自己的意志，从低温媒介流动到高温媒介。Clausius 的表述证明这一过程不是不可能的，如果将外电源引入到此过程，作为冷却器的一个条件，则可证明冷却机只有在压缩机被外接电源驱动的情况下才能工作。

2.2　热传递

热传递是由空间温差引起的热能传递。当温差存在于媒介或介质之间时，热传递便会发生。

热传递有 3 种不同的模式。静止的固体或液体媒介中存在温度梯度，这样的热传递称为传导。如果热传递发生在存在温差的流动液体及其表面之间，被称为对流。热传递的第 3 种模式是热辐射或简单辐射，指的是不同温度的无介质的两个界面发生热传递。

2.2.1　传导

传导指的是由于温度梯度的存在能量在介质中转移。原子和分子活动促进传导过程。也可看成能量从物质中活泼的原子或分子中转移到相邻相对不活泼的原子或分子中。

认为物质中存在温度梯度并假设无大量或宏观运动。物质任意一点的温度与在相似温度下原子或分子的能量相关。这种能量分子和原子的无规则位移运动，与内旋转和振动运动有关。高温意味着高能量。当相邻原子或分子发生碰撞时，能量就会从活泼的转移到相对不活泼的原子或分子上。此外，由于无规则运动，电子、原子或分子从上到下通过一个假设平面，上部比下部的颗粒有更高的能量。净能量是从高温地方向低温地方转移，如图 2.1 所示。颗粒间碰撞提高能量传递。

对于热传导过程，传热速率取决于傅里叶定律。傅里叶定律阐述了某一方向上的传热速率与该方向上的温度梯度、吸热方向上的横截面积区域的热流量成正比。对于一维的传热模式，傅里叶定律可以表述为

$$q_x = -k\frac{\mathrm{d}T}{\mathrm{d}x} \tag{2.6}$$

式中，q_x 是在 x 方向上垂直于单位横截面积的传热速率（$\mathrm{Wm^{-2}}$），称为热通量，热通量正比于其对应方向上的温度梯度 $\mathrm{d}T/\mathrm{d}x$；k 是热导率（$\mathrm{Wm^{-1}\cdot K}$），描述基体材料的性质；负号表示传热方向与温度下降方向一致。

已知横截面积为 A，傅里叶定律也可以写成

$$\boldsymbol{Q}_x = -kA\frac{\mathrm{d}T}{\mathrm{d}x} \tag{2.7}$$

式中，$\boldsymbol{Q}_x$ 是热通量（W），表示在 x 方向上通过横截面积 A 的热流量的传热速率。

热是一种能量，单位是焦耳（J）。因此，传热量是焦耳每秒（Js^{-1}）。这可以更直接地表示热传递这个概念。一些文献也用瓦特作为传热速率的单位来描述功率这个概念。功率就是做功的速率或能量转换速率。采用这个概念可以与热力学第一定律更紧密地联系起来。

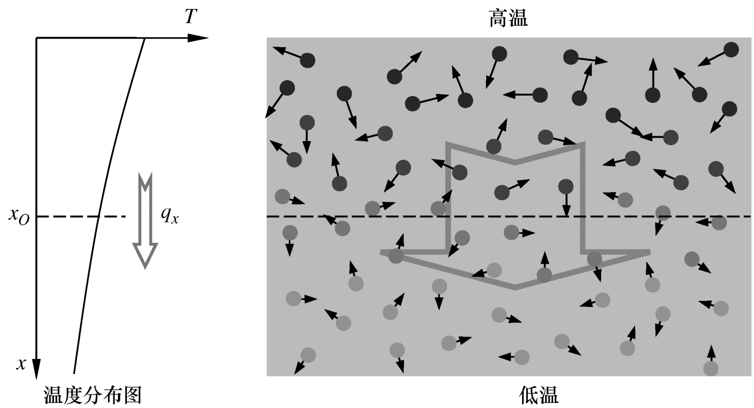


图 2.1　传导传热机制

傅里叶定律定量地描述了热通量或热流量。热流的方向垂直于传热的横截面。换言之，热流的方向总是垂直于恒定的温度表面。因此，热流也可以认为是一种矢量。对傅里叶定律更为概括地陈述的传导方程如下：

$$q_x = -k\left(\boldsymbol{i}\frac{\partial T}{\partial x}+\boldsymbol{j}\frac{\partial T}{\partial y}+\boldsymbol{k}\frac{\partial T}{\partial z}\right) \tag{2.8}$$

傅里叶定律是现象的或经验的。它不是从第一定律而是从观察现象发展出来的。热传导是材料重要的性质，傅里叶定律被认为是热传导的基石。它适用于所有材料、所有形态，不管是固态、液态还是气态。

2.2.2　对流

能量传递除了通过原子或分子水平的无规则运动之外，也可以通过流体大量或宏观运动传递能量。在存在温度梯度时，这种运动有助于热传递。总的热传递速率是分子的无规则运动及大量流体运动导致热传递的总和。我们通常用对流来描述这类热传递。另一个术语平流指的是仅因为大量流体运动导致的热传递。

在化学工程中，当流体和固体表面温度不同且流体流过固体表面时就会发生对流热传递。如果流体流动是由外力引起，如泵或风扇促使流体流过固体表面，这种热传递被称为强制对流。相反，如果流体流动是因为温差、浓度差引起的浮力导致的，这类热传递称为自由对流。

由于固体—流体相互作用导致流体流过固体表面，出现表面流速从零变到限定流速的大量流体的区域，该区域称为流体动力层或流速边界层。同样，如果在固体

表面和流体间存在温差，该区域称为热边界层，其存在于流体中。流体表面和内部的温差的变化发生在热边界层。

如图 2.2 所示，在固体表面和流体间的界面上，流体流速为零则热只能通过分子运动传递。在接近固体表面的边界层流体流速低，分子无规则运动限制了热传递。边界层随流体沿流动方向推移而增大，因此，流体运动的传热速率也随之增大。该层会发生热传递并沿着流动的方向传输，最终流出边界层，流到体相流体中。

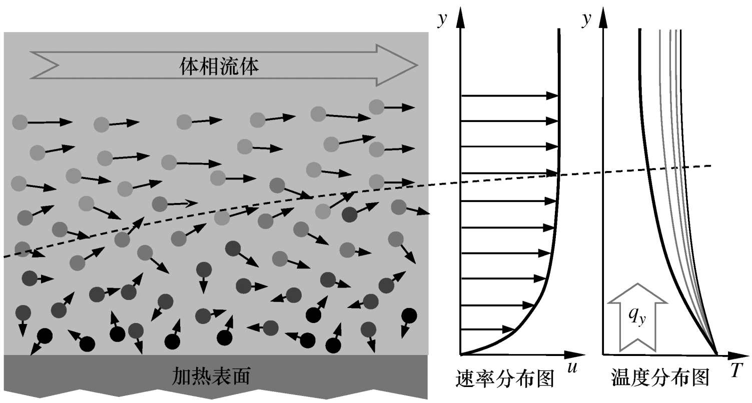


图 2.2　对流热传递和边界层

热对流是一个相对复杂的过程。在对流过程中，流体扮演了非常重要的角色。温度为 T_w 的固体表面及温度为 T_f 的流体流过其表面产生的对流热传递可以表述为

$$q = h(T_f - T_w) \tag{2.9}$$

式中，q 为对流热通量（Wm^{-2}），与固体表面和流体的温差成比例；h 是对流热传递系数（$Wm^{-2} \cdot K$）。

式（2.9）就是牛顿冷却定律。

对流热传递系数 h 取决于边界层的条件。它随流动类型的不同（即片层流动或扰动）、固体结构的几何形态、流体的物理性质、温度和固体表面所处位置的不同而发生改变，其也取决于强制对流或自由对流的类型。在工程应用中，h 在固体表面的位置的变化，考虑的就是平均值 h_m 而不是局部值 h。流体流过物体会形成几何图形，如平盘或圆形管道，热传递系数可以测定分析。对于复杂的几何图形，实验方法常被用于测定分析平均热传递系数 h_m。

2.2.3　辐射

如图 2.3 所示，由于所有物质都不是处于绝对零度㊀，所以它们会持续不断地

㊀　绝对零度，0K 或 −273.15℃。

放出能量。所发射的能量称为热辐射，其来自于物质表面所弹出的热能。其放射的辐射能通过电磁波传递。通过辐射进行的能量传递不需要表面存在介质，但是传导和对流需要。物体上任意一点都可以发射或吸收所有方向的辐射。

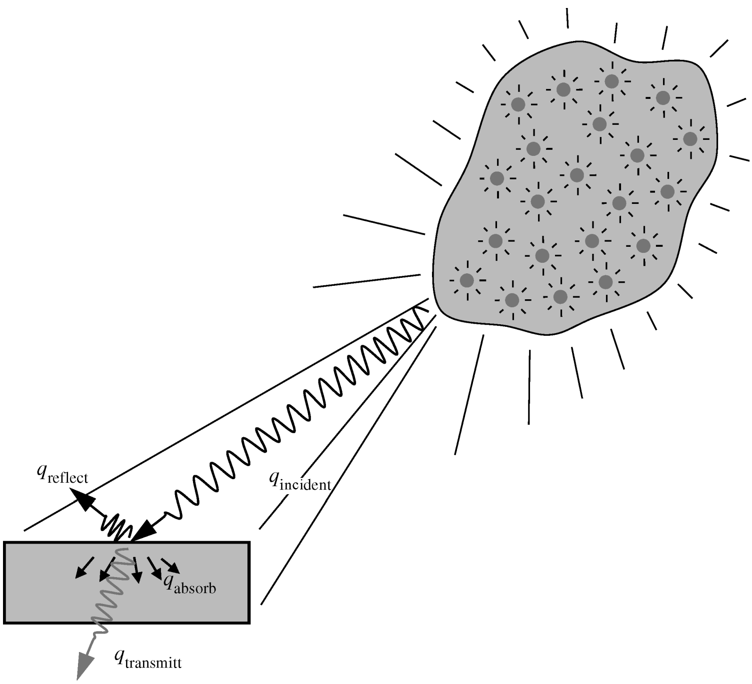


图 2.3 辐射热交换

单位面积释放出的辐射能的速率称为表面发射功率。斯特藩-波尔兹曼（Stefan-Boltzmann）定律影响所发射的最大辐射通量：

$$E_b = \sigma T_s^4 \tag{2.10}$$

式中，T_s 是物体表面绝对温度（K）；σ 是斯特藩-波尔兹曼常数（$\sigma = 5.6697 \times 10^{-8} W (m^2 \cdot K^4)^{-1}$）；$E_b$ 是黑体辐射功率。只有理想辐射体或所谓黑体才能发射最大辐射通量。

真实表面发射的辐射通量小于黑体发射功率，E_b 由下式得到：

$$q = \varepsilon \sigma T_s^4 \tag{2.11}$$

式中，ε 是表面的辐射特性，正如放射率变化范围是从零到单位值。

这一性质显示表面放射能效率与黑体有关。材料表面粗糙度对放射率有很大影响。

当辐射通量 q_{inc} 入射到黑体内时，它会完全被黑体吸收。但是，如果辐射通量 q_{inc} 被非黑体吸收，如果它是半透明的，则一部分被表面反射，一部分被吸收，一部分被发射。每一部分的比率分别被称为反射率、吸收率和发射率。能量被物体吸

收的关系式为

$$q_{abs} = \alpha q_{inc} \tag{2.12}$$

根据基尔霍夫定律的辐射状态可知吸收率与发射率相等。

当不同温度的两个物体相互可“见”时，可通过辐射而互相交换热。众所周知，只有一部分辐射能离开一个表面到达另外一个，剩下的损失在周围环境中。形态系数的概念用来描述一部分辐射能离开一个表面直接冲到另一个表面。两表面间的净热传递可写成

$$Q = A_1 F_{A_1-A_2}\sigma T_1^4 - A_2 F_{A_2-A_1}\sigma T_2^4 \tag{2.13}$$

式中，$F_{A_1-A_2}$和 $F_{A_2-A_1}$分别是从 A_1 到 A_2 表面的形态系数和从 A_2 到 A_1 表面的形态系数。

2.3　质量传递

质量传递指的是由浓度梯度引起的混合物中的扩散运动。

如果混合物是不均匀的，一定区域内某一成分的分子数高于其他成分。分子由于无规则运动碰撞几率增大，这种碰撞会逐渐扩散到整个混合物。这种扩散称为常扩散。另外，压力梯度、热梯度和外力也会引起物质运动。这些附加效应也被称为扩散。

不同种物质的浓度也有多种表示方法。通常，成分用质量浓度表示，ρ_i（kgm^{-3}），每单位体积溶液中 i 种物质的质量；以及摩尔浓度，c_i（$molm^{-3}$），每单位体积溶液中 i 种物质的摩尔数。我们也利用质量分数，$\omega_i = \rho_i/\rho$，定义为溶液的总质量密度除以质量浓度；摩尔分数，$x_i = c_i/c$，溶液的总摩尔密度除以 i 种物质的摩尔浓度。

扩散质量传递的速率可由菲克（Fick）第一定律说明，这一过程可由图 2.4 所示。菲克定律指出给定方向上的质量传递速率与该方向上的浓度梯度成比例，是一维传递，菲克定律是

$$j_{Ax} = -D_{AB}\frac{dc_A}{dx} \tag{2.14}$$

式中，j_{Ax}表示 A 物质在 x 方向的摩尔质量通量除以垂直于传递方向的摩尔质量通量（$mol\ (s \cdot m^2)^{-1}$）；D_{AB}是扩散系数或 A 和 B 二元溶液中 A 物质的扩散率（m^2s^{-1}），扩散率是组成与周围环境的一种特性，可应用于气体、液体和固溶体；负号强调扩散的方向是浓度减小的方向。

知道横截面积 A，菲克定律也可表示为

$$J_{Ax} = -D_{AB}A\frac{dc_A}{dx} \tag{2.15}$$

式中，J_{Ax}是扩散质量流（$mols^{-1}$）。

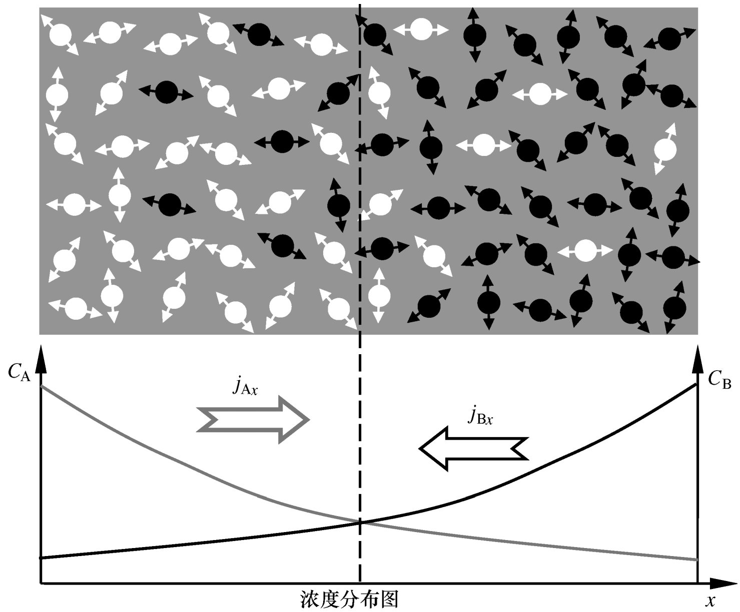


图 2.4　扩散质量传递

质量扩散率，D_{AB}，在二元系统中是温度、压力和组成成分的函数。大多数二元混合物中 D_{AB} 数据的有效性在范围和准确度上很有限。与 D_{AB} 相关的限定范围基于理论比试验多。

二元混合气体在低压下，D_{AB} 与压强成反比，随温度升高而增大，并且几乎与气体偶的组成无关。二元气体混合物低压扩散率可通过结合动力理论和对应状态论证估量。把混合气体设想为理想气体，扩散率可由下式表示：

$$D_{AB} = 0.0018583 \frac{\sqrt{T^3\left(\frac{1}{M_A}+\frac{1}{M_B}\right)}}{P\sigma_{AB}^2\Omega_{D,AB}} \tag{2.16}$$

式中，T 是温度（K）；P 是总压力（atm）[㊀]；M 是 A 和 B 的摩尔质量（$gmol^{-1}$）；σ_{AB} 是 Lennard-Jones 参数（Ångström）；$\Omega_{D,AB}$ 是温度的无量纲参数与 A 和 B 分子间势场。

富勒（Fuller）方程也被用来估量扩散率：

$$D_{AB} = \frac{1.013\times10^{-2}T^{1.75}\left(\frac{1}{M_A}+\frac{1}{M_B}\right)^{0.5}}{P\left[\left(\sum\nu_i\right)_A^{\frac{1}{3}}+\left(\sum\nu_i\right)_B^{\frac{1}{3}}\right]^2} \tag{2.17}$$

式中，$(\sum\nu_i)_A$ 和 $(\sum\nu_i)_B$ 分别是 A 和 B 的扩散体积，富勒相关式有较好的准

㊀ 1atm = 101.325kPa。——译者注

确性。

2.4　流体力学

流体定义为一个物质遭受一定数值的剪切应力而发生连续变形。术语中连续变形通常被认为是流动的。导致变形的力是剪切应力，它是通过具有流动特性的称之为黏度的传递。这类流动机制如图 2.5 所示。流体是物相的支组，包括气体、液体、离子体和塑胶固体。

流动也可把物质从一个地方运输到另一个地方，其运输机理是通过扩散质量传递而区分开，这是应力导致的宏观运动。

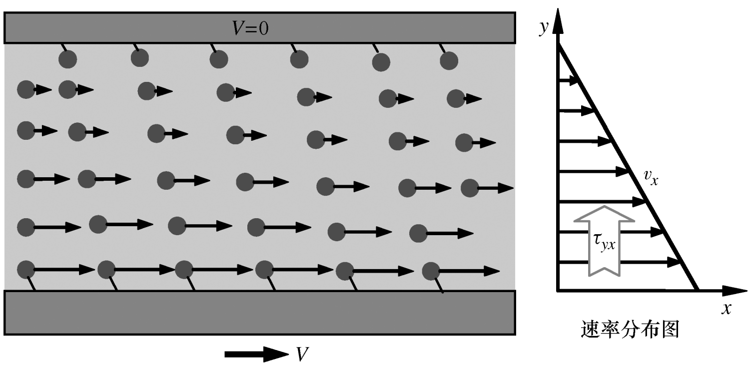


图 2.5　牛顿流体机制

2.4.1　黏性流

剪切力使每单位面积流体保持运动，如牛顿黏度定律，它与局部速度梯度的负数成比例：

$$\tau_{yx} = -\mu \frac{dv_x}{dy} \tag{2.18}$$

式中，τ_{yx}是 y 方向上的应力（kg（m · s^2）$^{-1}$），且力的方向平行于 x 方向；v_x 是朝 x 方向的速度（ms^{-1}）；μ 是流体黏度（kg（m · s）$^{-1}$）。

黏度是描述流体内部电阻的流动并可作为衡量流体摩擦的手段。黏度，分子结构角度认为，其来源于分子间的凝聚力。流体的黏度易于测试，可通过常用设备检测并且它的温度和压力的关系可通过手册和书中的注解及参考文献查到。

剪切力也可解释为动量，mv（kg · ms^{-1}），是运动物体的重要性质，它是物体运动的内在推动力，是运动的量度。如图设 $y=0$，在运动表面的相邻层，运动流体获得一定量的 x 动量，即在 x 方向上的动量，然后流体一部分动量分给相邻层使其运动仍保持在 x 方向。依次的，x 动量通过在 y 方向上的流体传递。因此，τ_{yx}也许可以理解为在 y 方向上的黏性 x 动量通量，或沿 y 方向的 x 动量传递通量。这个

解释更好地结合了动量运输过程的分子本性，并常用于分析流体机理。

2.4.2　速度分布

一个流动系统的速度剖面图可通过结合动量平衡方程和式（2.18）说明力影响液体黏度。困难在于通过控制流体体积建立动量平衡。流动的流体在稳态下的动量平衡方程式如下：

［流入动量速率］-［流出动量速率］+［总的作用在系统上的力］=0（2.19）

这个方程可用于通过限制体积而改进溶液剪切力 τ。式（2.18）可用于替换 τ。解决了给定边界条件产物不同的方程就会解决不同位置的 v。

2.4.3　伯努利方程

先前的探索者例如丹尼尔-伯努利（Daniel Bernoulli）展开相关描述流动流体的能量、速度和压力的宏观关系。相关式广为流传并被广泛应用于伯努利原理。

伯努利原理指出对于无黏性流动，流体流动的速度增加与压力的减小或流动的势能减小同时发生。大部分液体和气体在低马赫数（Mach number）有低黏度并可近似认为是零黏度或无黏性流动，若其中只有宏观或明显的性质，如压力、温度和能量，则被考虑。对于不可压缩的流动，单看伯努利原理，在沿流线的任意一点都有效，可通下式表达：

$$P + \frac{1}{2}\rho v^2 + \rho gh = \text{常数} \tag{2.20}$$

或

$$P_A + \frac{1}{2}\rho v_A^2 + \rho gh_A = P_B + \frac{1}{2}\rho v_B^2 + \rho gh_B \tag{2.21}$$

式中，P 是压强（Pa）；v 是流体速度（ms^{-1}）；ρ 是流体密度（kgm^{-3}）；g 是重力常数（kgs^{-2}）；h 是参考面上的点到参考面的距离（m）；下标 A 和 B 代表在流线上的任意两点。

实质上，伯努利原理结合了能量守恒和流体动量。它说明，在稳流中，在流线上所有流体机械能相等。例如，如果流体从蓄水池中向外流，在流线上所有形式的能量都相等，因为在一个蓄水池中，每单位质量的能量（压强和重力势 ρgh 总和）在每一点都是常数。

除了速度和压力分布，性能研究或课题，例如层流或乱流，压力作用和外力，液压，在导管中的黏流，在明渠中流动，通过潜水体流动，压缩流动，空气动力和多相流动，对于应用有重要影响。先前的课题大部分已经研究得很深入了，无论是实验上还是理论上。尤其，有了计算流体动力学（CFD）的帮助，这是基于牛顿黏度第一定律，改变方程式和数字计算，这一先进课题使研究更为容易。

2.5　变化方程式

有关热传导的傅里叶定律、质量扩散的菲克定律以及动量传递的牛顿定律描述

了 3 种基本传输过程的能量来源。许多稳态性质都来自于这些基本方程，如稳态温度分布、浓度分布和速度分布等情况。但是，仅通过公式计算不能完全描述出稳态的传输过程，还需要了解这些传输过程的动力学情况和转变速率。根据质量守恒定律、动量守恒定律和能量守恒定律可以分析这些动力学问题。

质量守恒方程和动量守恒方程可以描述热能和热传输不受周围环境影响（恒温）下纯流体的黏度流动问题。单一流体的流动会和周围环境交换能量，复合流体的混合物则需要通过能量守恒和质量守恒方程来弄清所包含的化学品种类。这些守恒方程有时也称为变化方程，正如它们所描述的在系统中不同时间、位置的速度变化、温度变化和浓度变化。

2.5.1　连续性方程

连续性方程是通过质量守恒定律来计算流动的流体中小体积元素而发展出来的。在稳态下，小体积元素的质量平衡公式为

$$[\text{质量聚集速率}] = [\text{流入质量速率}] - [\text{流出质量速率}] \tag{2.22}$$

质量聚集速率是 $\Delta V\dfrac{\partial\rho}{\partial t}$，在可控容积中质量流入或流出速率分别为 $\Delta A_x\rho v_x$、$\Delta A_y\rho v_y$ 和 $\Delta A_z\rho v_z$，这里 V 和 A 分别是不同元素的体积和面积。

如果 ΔV 接近于 0，式（2.22）变为

$$\frac{\partial\rho}{\partial t} = -\left(\frac{\partial}{\partial x}\rho v_x + \frac{\partial}{\partial y}\rho v_y + \frac{\partial}{\partial z}\rho v_z\right) \tag{2.23}$$

该方程描述了质量速度矢量 $\rho \boldsymbol{v}$ 的固定点的密度改变速率。

2.5.2　运动方程

运动方程是从小体积物质的能量守恒发展而来的：

$$\begin{aligned}[\text{动量聚集速率}] = {} & [\text{动量流入速率}] - [\text{动量流出速率}] \\ & + [\text{总的作用在系统上的力}]\end{aligned} \tag{2.24}$$

因为动量是一个矢量，可以直接将其放入一个正交坐标系中，即 x 轴、y 轴和 z 轴，再分别考虑 x 方向、y 方向和 z 方向的动量平衡。

在 x、y 和 z 三个方向上的动量聚集速度分别是 $\Delta V\dfrac{\partial}{\partial t}\rho v_x$、$\Delta V\dfrac{\partial}{\partial t}\rho v_y$ 和 $\Delta V\dfrac{\partial}{\partial t}\rho v_z$。

动量流出和动量流入包括两部分：一部分来自流体流动；另一部分来自速度梯度。需要注意的是，在某一方向上的动量成分，如 x 动量，可以从 3 个方向参与到可控容积内。x 方向上，动量流入和流出速率分别是 $\Delta A_x\rho v_x v_x$、$\Delta A_y\rho v_y v_x$ 和 $\Delta A_z\rho v_z v_x$。

同样，由于速率梯度引起可控或不可控的容积内，x 方向动量速率为 $\Delta A_x\boldsymbol{\tau}_{xx}$、$\Delta A_y\boldsymbol{\tau}_{yx}$ 和 $\Delta A_z\boldsymbol{\tau}_{zx}$。

大多数情况下，最重要的力是由流体压力 p 和单位质量所受重力 g 引起的。所以 x 方向的受力总和为 $\Delta A_x\Delta p_x + \Delta V\rho g_x$。

除上述动量平衡方程外，我们得到了 3 个方面的动量成分的 3 个运动方程。

x 成分的运动方程：

$$\begin{aligned}\frac{\partial}{\partial t}\rho v_x &= -\left(\frac{\partial}{\partial x}\rho v_x v_x + \frac{\partial}{\partial y}\rho v_y v_x + \frac{\partial}{\partial z}\rho v_z v_x\right) \\ &\quad -\left(\frac{\partial}{\partial x}\tau_{xx} + \frac{\partial}{\partial y}\tau_{yx} + \frac{\partial}{\partial z}\tau_{zx}\right) - \frac{\partial p}{\partial x} + \rho g_x\end{aligned} \tag{2.25}$$

y 成分的运动方程：

$$\begin{aligned}\frac{\partial}{\partial t}\rho v_y &= -\left(\frac{\partial}{\partial x}\rho v_x v_y + \frac{\partial}{\partial y}\rho v_y v_y + \frac{\partial}{\partial z}\rho v_z v_y\right) \\ &\quad -\left(\frac{\partial}{\partial x}\tau_{xy} + \frac{\partial}{\partial y}\tau_{yy} + \frac{\partial}{\partial z}\tau_{zy}\right) - \frac{\partial p}{\partial y} + \rho g_y\end{aligned} \tag{2.26}$$

z 成分的运动方程：

$$\begin{aligned}\frac{\partial}{\partial t}\rho v_z &= -\left(\frac{\partial}{\partial x}\rho v_x v_z + \frac{\partial}{\partial y}\rho v_y v_z + \frac{\partial}{\partial z}\rho v_z v_z\right) \\ &\quad -\left(\frac{\partial}{\partial x}\tau_{xz} + \frac{\partial}{\partial y}\tau_{yz} + \frac{\partial}{\partial z}\tau_{zz}\right) - \frac{\partial p}{\partial z} + \rho g_z\end{aligned} \tag{2.27}$$

根据速度梯度和流体性质所得到的不同压力下的运动方程，连同连续方程、状态方程 $p = p(\rho)$、黏度随密度发生改变的方程 $\mu = \mu(\rho)$ 及边界和初始条件决定了等温条件下的流体流动的压力、密度和不同成分的流速。

对于常数 ρ 和 μ，这些运动方程可以综合在一起简化为

$$\rho\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} = -\nabla p + \mu\nabla^2\boldsymbol{v} + \rho\boldsymbol{g} \tag{2.28}$$

该方程是 Navier-Stokes 方程。

2.5.3 能量平衡方程

流体带有能量。在确定条件下，一种形式的能量会转换成其他形式。比如，流体的潜能会转变为动能。流体也可能会与周围环境交换能量。想要完整地描述流体的性质，需要利用能量平衡方程、连续性方程以及运动方程来计算流体的能量守恒。

小体积流体的能量平衡可以描述为

[内能和动能的聚集速率] = [内能和动能的转入速率] − [内能和动能的转出速率] + [热能交换速率] − [系统对环境做功的速率]　　(2.29)

小体积的内能和动能的聚集速率为 $\Delta V\frac{\partial}{\partial t}\left(\rho U + \frac{1}{2}\rho v^2\right)$。这里，$U$ 是单位质量流体内能，v 是局部流体速度。

内能和动能的转入与转出速率是 $\Delta A_x v_x\left(\rho U + \frac{1}{2}\rho v^2\right)$、$\Delta A_y v_y\left(\rho U + \frac{1}{2}\rho v^2\right)$和

$\Delta A_z v_z\left(\rho U+\frac{1}{2}\rho v^2\right)$。

热能转入或转出速率是 $\Delta A_x q_x$、$\Delta A_y q_y$ 和 $\Delta A_z q_z$。

克服重力做功速率是 $-\rho\Delta V v_x g_x$、$-\rho\Delta V v_y g_y$ 和 $-\rho\Delta V v_z g_z$。

针对静态压力做功的速率是 $\Delta A_x p v_x$、$\Delta A_y p v_y$ 和 $\Delta A_z p v_z$。

最后，为补偿黏滞力做功速率是 $\Delta A_x(\tau_{xx}v_x+\tau_{xy}v_y+\tau_{xz}v_z)$、$\Delta A_y(\tau_{yx}v_x+\tau_{yy}v_y+\tau_{yz}v_z)$ 和 $\Delta A_z(\tau_{zx}v_x+\tau_{zy}v_y+\tau_{zz}v_z)$。

因此，能量平衡方程的基本形式是

$$\begin{aligned}\frac{\partial}{\partial t}(\rho U+\frac{1}{2}\rho v^2) = &-\left(\frac{\partial}{\partial x}v_x(\rho U+\frac{1}{2}\rho v^2)+\frac{\partial}{\partial y}v_y(\rho U+\frac{1}{2}\rho v^2)+\frac{\partial}{\partial z}v_z(\rho U+\frac{1}{2}\rho v^2)\right)\\ &-\left(\frac{\partial q_x}{\partial x}+\frac{\partial q_y}{\partial y}+\frac{\partial q_z}{\partial z}\right)+\rho(v_x g_x+v_y g_y+v_z g_z)\\ &-\left(\frac{\partial}{\partial x}pv_x+\frac{\partial}{\partial y}pv_y+\frac{\partial}{\partial z}pv_z\right)\\ &-\left(\frac{\partial}{\partial x}(\tau_{xx}v_x+\tau_{xy}v_y+\tau_{xz}v_z)+\frac{\partial}{\partial y}(\tau_{yx}v_x+\tau_{yy}v_y+\tau_{yz}v_z)\right.\\ &\left.+\frac{\partial}{\partial z}(\tau_{zx}v_x+\tau_{zy}v_y+\tau_{zz}v_z)\right)\end{aligned} \tag{2.30}$$

正如式（2.30）所示，能量平衡方程可以分为两部分：热能和动能。分离成分与式（2.30）的动能相关得到热能方程如下：

$$\rho\frac{\mathrm{D}U}{\mathrm{D}t}=|-(\boldsymbol{\nabla}\cdot\boldsymbol{q})-p(\boldsymbol{\nabla}\cdot\boldsymbol{v})-(\boldsymbol{\tau}:\boldsymbol{\nabla v}) \tag{2.31}$$

动能方程是

$$\rho\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}v^2\right)=-p(\boldsymbol{\nabla}\cdot\boldsymbol{v})-(\boldsymbol{\nabla}\cdot p\boldsymbol{v})+p(\boldsymbol{v}\cdot\boldsymbol{g})-(\boldsymbol{\nabla}\cdot[\boldsymbol{\tau}\cdot\boldsymbol{v}])+(\boldsymbol{\tau}:\boldsymbol{\nabla v}) \tag{2.32}$$

对于大部分工程应用，用热能方程分析流体温度和热熔比内能更方便：

$$\rho C_v\frac{\mathrm{D}T}{\mathrm{D}t}=-(\boldsymbol{\nabla}\cdot\boldsymbol{q})-T\left(\frac{\partial p}{\partial T}\right)_V(\boldsymbol{\nabla}\cdot\boldsymbol{v})-(\boldsymbol{\tau}:\boldsymbol{\nabla v}) \tag{2.33}$$

此方程是分析流体温度 T 的能量方程，可用于计算反应温度及其温度分布情况。

2.5.4　连续性方程的种类

对于多组分流体，通过反应，每种连续性方程都可能以相似的途径被改进而形成通用的连续性方程：

$$\frac{\partial c_{\mathrm{A}}}{\partial t}=-\left(\frac{\partial}{\partial x}c_{\mathrm{A}}v_x+\frac{\partial}{\partial y}c_{\mathrm{A}}v_y+\frac{\partial}{\partial z}c_{\mathrm{A}}v_z\right)+r_{\mathrm{A}} \tag{2.34}$$

式中，r_{A} 是依据浓度进行的化学反应产生 A 的速率，这是物质 A 的连续性方程。

同样的，如果化学反应吸收或释放热，能量方程也需要包括在热反应等方面的更新。

2.6 化学反应

化学反应是一个过程，在一种或多种化学种类中检测分子数恒等式无意义并假设一个新形式，通过改变化学组成的种类或原子数或通过改变这些原子的结构或形态。单化学种类的恒等式决定于这些原子的种类、数量和形态。众所周知，一个化学反应发生的地方，所有参与反应的物质的总质量保持不变。但是，当考虑到独立物质涉入反应中，特定种类的质量消失速率或形成速率可以被考虑进来。

2.6.1 反应速率

化学反应进行的速率可以表示为反应物的消失速率或产物的形成速率。显而易见的是，反应速率 r_j(mol(m^3 · s)$^{-1}$) 被定义为物质 j 在单位时间单位体积消失的摩尔数。

考虑一个反应：

$$aA + bB \rightarrow cC + dD$$

每种物质的反应速率定义为

$$\begin{aligned} -r_A &= -\frac{1}{\Delta V}\frac{dN_A}{dt} \\ -r_B &= -\frac{1}{\Delta V}\frac{dN_B}{dt} \\ r_C &= -\frac{1}{\Delta V}\frac{dN_C}{dt} \\ r_D &= -\frac{1}{\Delta V}\frac{dN_D}{dt} \end{aligned} \tag{2.35}$$

式中，N_i 是物质 i 在控制体积的摩尔数及 ΔV 时控制体积中的体积。

在一个反应中，理想比系数 a、b、c 和 d 不是必须相等的，并且每种物质的反应速率也不相等，但它们需要满足以下关系：

$$ar_A = br_B = cr_C = dr_D \tag{2.36}$$

为避免意思模糊，速率被普遍定义为

$$r = \frac{1}{v_j \Delta V}\frac{dN_j}{dt} \tag{2.37}$$

式中，v_j 是物质 j 的化学计量系数。

当反应不发生体积变化时，例如，反应发生在液相中，反应速率可依浓度而定：

$$-r_j = \frac{dC_j}{dt} \tag{2.38}$$

学习反应速率的两个主要目的是确定需要的反应器尺寸和计算反应器中反应物的成分。大部分工业反应器可用以下 4 个很理想的反应器模型：批式反应器、连续搅拌槽反应器（CSTR）、活塞流反应器（PFR）和填充床反应器。知道了反应速率和反应器模型，反应器尺寸就可以确定了。转换率参数 x_j 的引入可方便地计算在反应器中反应物组分及反应进行的程度。转换率参数 x_j 是物质 j 已经参加反应的摩尔数除以物质 j 投入系统的摩尔数：

$$X_j = \frac{\text{物质 } j \text{ 已经参加反应的摩尔数}}{\text{物质 } j \text{ 投入系统的摩尔数}} \tag{2.39}$$

所以，可以根据转换率表示反应速率 X_j：

$$-r_j = \frac{1}{\Delta V} \cdot N_{j0}\frac{\mathrm{d}X_j}{\mathrm{d}t} \tag{2.40}$$

式中，N_{j0}是 j 进料的初始摩尔数。

反应速度取决于浓度和温度。反应动力学实验结果显示，反应速度与反应物的浓度函数成比例。例如，物质 A 的反应速率可表达为

$$-r_{\mathrm{A}} = k_{\mathrm{A}} C_{\mathrm{A}}^{\alpha} C_{\mathrm{B}}^{\beta} \tag{2.41}$$

式中，k_{A} 是速率常数；α 和 β 是反应级数。

这些参数可通过实验得到。

2.6.2　可逆反应

理论上，几乎所有的反应皆可逆，反应表达为

$$a\mathrm{A} + b\mathrm{B} \rightleftharpoons c\mathrm{C} + d\mathrm{D}$$

当 A 与 B 反应产生 C 和 D 的速率与 C 与 D 反应产生 A 和 B 的速率相等时，则反应达到平衡。在平衡状态下，物质的浓度不再变化且看起来反应停止了。

对于可逆反应，速率可表达为

$$r = kC_{\mathrm{A}}^{\alpha} C_{\mathrm{B}}^{\beta} - k' C_{\mathrm{C}}^{\gamma} C_{\mathrm{D}}^{\delta} \tag{2.42}$$

式（2.42）右边首项是正反应速率，其中 k 是正速率常数。右边第二项是逆反应常数，k'是逆反应速率常数。

在反应初期，当产物浓度低时，逆反应通常无足轻重。随着产物浓度增加，逆反应速率逐渐增加直到反应达到平衡。然后每种物质的浓度停止变化。

在平衡状态下，正向和逆向反应速率相等：

$$kC_{\mathrm{A}}^{\alpha} C_{\mathrm{B}}^{\beta} = k' C_{\mathrm{C}}^{\gamma} C_{\mathrm{D}}^{\delta} \tag{2.43}$$

因此，通过以下方程可以推导出反应平衡常数：

$$K_{\mathrm{c}} = \frac{k}{k'} = \frac{C_{\mathrm{C}}^{\gamma} C_{\mathrm{D}}^{\delta}}{C_{\mathrm{A}}^{\alpha} C_{\mathrm{B}}^{\beta}} \tag{2.44}$$

平衡常数取决于温度并可通过热力学和物理化学计算出来。对于放热反应，平衡常数随温度升高而减小，吸热反应中，其随温度升高而增大。

2.6.3　反应热

每种化合物涉及化学反应都会生成焓。每种物质的形成都要求能量与生成焓相

等，然而每种物质消耗和释放出的能量数值相等。这意味着在化学反应发生过程中，热会被释放或吸收，所有产物的生成焓与反应物的不总是一致。

生成焓 $H(\mathrm{Jmol}^{-1})$，是一个元素或化合物在参考状态下的基本性质，它可被看成物质由化学组成在特定状态下的焓。

因此，反应热是

$$Q = cH_C + dH_D - aH_A - bH_B \tag{2.45}$$

负结果表示产物的生成焓比反应物的少，并且额外的化学能作为反应热被释放到周围环境中，这种反应被称为放热反应。正结果意味着反应从周围吸收热量来保证反应进行，这种反应被称为吸热反应。

反应热在计算流动的反应流能量中扮演重要角色。如果在系统中进行质量守恒计算，由浓度决定物种可用于计算涉及反应的热与式（2.45）计算反应热相似。

2.7 注解和参考文献

本章节提供了关于热力学、热和质量传递、流体力学和反应动力学的一些观点，主要基于以下文献：

Cengel and Boles（1994）and Smith and Ness（1987a）for thermodynamics；Incropera *et al.*（2006）and OziSik（1985）for heat transfer；Treybal（1980）for mass transfer；Munson *et al.*（2009）for fluid mechanics；Bird *et al.*（1960）and Welty *et al.*（1984）for transport process；and Fogler（1999），Levenspiel（1999）and Ulrich（1993）for chemical reaction.

若想更深入、更宽泛地讨论关于本章的内容，读者可以参考原始教科书了解更多知识。

第 3 章　系统辨识 I

3.1　离散时间系统

系统辨识是一种有着广泛应用的技术，通常作为一种方法用于控制系统设计以获得动态模型。系统辨识是一种最具挑战性和耗时的任务，这通常是被有经验的工程师在加工工业的控制系统的实施工作中认可的。本章的目标和接下来要讲的是让读者对系统辨识的基本原则和先进原理有所了解，这将在本书后面的章节中用到。

在前面的章节中，动态建模被认为是源自第一性原理的。第一性原理模型通常表示为非线性和/或偏常微分方程。这些模型对于高保真动态模拟和过程分析是有用的。简化第一性原理模型，例如通常包含了集中的非线性常微分方程，经常用于控制器的设计是典型的非线性控制器设计。进一步简化模型，如那些由线性常微分方程组成，使当初非线性的形式线性化，是最常见的动态模型，用于传统控制策略的设计，例如 PID 和 MPC。由于在实践中最常用到的是线性动态模型，一个非常现实的问题是这些线性模型是否可以直接从实验数据的分析中开发出来。这个问题会在本章和后面的章节得到解释。本章主要介绍系统辨识的一些基本知识，包括系统的离散时间表示、信号的获得或设计，这些都是为了系统辨识的实验和模型结构且将会用于系统辨识。通过从第一性原理推导出常微分方程来开始本章，正如在前面的章节中介绍的开发离散时间模型那样。

解决常微分方程数值的最简单的方法是通过有限差分法，其中倒数运算经由一个差分近似操作：

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} \approx \frac{y_t - y_{t-1}}{T_s} \tag{3.1}$$

式中，T_s 为采样间隔；y_t 为 $y(t)$ 的采样数量。

同样，一个二阶导数的近似值可表示为

$$\frac{\mathrm{d}^2 y(t)}{\mathrm{d}t^2} \approx \frac{y_t - 2y_{t-1} + y_{t-2}}{T_s^2}$$

一阶差分方程具有以下形式：

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = f(y(t), u(t)) \tag{3.2}$$

把式（3.1）带入式（3.2）估算函数 $f(y(t), u(t))$，y 和 u 的值在 $t-1$ 时刻（即 y_{t-1} 和 u_{t-1}）得到：

$$\frac{y_t - y_{t-1}}{T_s} \approx f(y_{t-1}, u_{t-1}) \tag{3.3}$$

式（3.3）可以写为

$$y_t \approx y_{t-1} + T_s f(y_{t-1}, u_{t-1}) \tag{3.4}$$

由于差分操作已被用于近似式（3.2），所得式（3.4）被称为差分方程。差分方程的阶数是由 y 的最高指数减去最低指数确定的。在这种情况下，该命令可以被确定为 $t-(t-1)=1$，因此，式（3.4）是一阶差分方程。差分方程可以用来预测 t 时刻 y 的响应，根据易得到的前一时刻 $t-1$ 的信息，例如 y_{t-1}和 $f(y_{t-1},\ u_{t-1})$。同样的，用所预测的 y_t 和给定的 u_t，y_{t+1}也可以预测。因此，这种类型的表达也就是所谓的递推关系。但是，由于差分方程是原微分方程的近似，所以所得到的预测 $\{y_t,\ t=1,\ 2,\ 3,\ \cdots\}$ 变得不太准确，这是由于 T_s 的增加。为了更好地逼近，选择更小的 T_s。但是，对于极小的 T_s 值，计算机舍入误差可能是错误的主要来源。

对于由线性微分方程描述的系统，另一个离散时间模型可以基于一个分段常数输入的解析解得到。如果相邻的时刻之间的输入变量实际上是恒定的，这种方法产生一个精确的离散时间模型。因此，这种分析方法消除了离散误差，固有的有限差分近似为重要的实际情况，数字计算机输出（过程输入）保持恒定的采样间隔。如果 D-A 转换设备作为一个零阶保持，确实是这样的。下面是确切的离散化的一个例子。

考虑到一个连续时间状态空间模型：

$$\dot{x}(t) = \boldsymbol{A}x(t) + \boldsymbol{B}u(t) \tag{3.5}$$

$$y(t) = \boldsymbol{C}x(t) \tag{3.6}$$

式中，$\boldsymbol{A}$、$\boldsymbol{B}$ 和 $\boldsymbol{C}$ 是常数矩阵。仔细考虑系统采样的离散时间 $t_i(i=1,\ 2,\ \cdots,)$ 和 $t_{i+1}-t_i=T_s$，输入是一个分段常数，即 $u(t)=u(t_i)$，$t_i \leqslant t < t_{i+1}$。积分式（3.5）在 $t=t_{i+1}$时解得 $x(t)$：

$$\begin{aligned} x(t_{i+1}) &= e^{A(t_{i+1}-t_i)}x(t_i) + \left[\int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-\xi)} B\mathrm{d}\xi\right]u(t_i) \\ &= e^{AT_s}x(t_i) + \left[\int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-\xi)} B\mathrm{d}\xi\right]u(t_i) \end{aligned}$$

定义 $\gamma = t_{i+1} - \xi$，然后

$$\begin{aligned} x(t_{i+1}) &= e^{AT_s}x(t_i) + \left[\int_{t_{i+1}-t_i}^{0} e^{A\gamma} B\mathrm{d}(-\gamma)\right]u(t_i) \\ &= e^{AT_s}x(t_i) + \left[\int_{0}^{T_s} e^{A\gamma} B\mathrm{d}\gamma\right]u(t_i) \end{aligned}$$

其结果是，在离散时间状态空间模型可以表示为

$$x(t_{i+1}) = e^{AT_s}x(t_i) + \left[\int_{0}^{T_s} e^{A\gamma} B\mathrm{d}\gamma\right]u(t_i)$$

$$y(t_i) = Cx(t_i)$$

使用速记符号（例如 x（t_i）被 x_i 取代），离散时间状态空间模型可以简化为

$$x_{i+1} = e^{AT_s}x_i + \left[\int_0^{T_s} e^{A\gamma}B d\gamma\right]u_i$$

$$y_i = Cx_i$$

在本书中，我们用离散时间为 t 的下角。所以我们用 t 替换 i 给出

$$x_{t+1} = e^{AT_s}x_t + \left[\int_0^{T_s} e^{A\gamma}B d\gamma\right]u_t \tag{3.7}$$

$$y_t = Cx_t \tag{3.8}$$

式（3.8）的推导没有近似被得出，这离散时间状态空间模型的解决方案是完全一样的连续时间状态空间模型的解，即式（3.5）和式（3.6），在采样时刻。出于这个原因，离散的这种做法被称为确切的离散。

同样，可以通过一个连续时间传递函数得到离散时间表示。考虑一阶传递函数，例如：

$$\frac{y(s)}{u(s)} = \frac{k}{\tau s + 1}$$

式中，$u(t)$ 是分段恒定信号，它可以被转换为一个差分等式：

$$\frac{dy(t)}{dt} = -\frac{1}{\tau}y(t) + \frac{k}{\tau}u(t)$$

式中，$u(t)$ 是分段恒定信号。定义一个状态变量为 $x(t) = y(t)$，空间状态表示可以被写为

$$\dot{x}(t) = -\frac{1}{\tau}x(t) + \frac{k}{\tau}u(t) \tag{3.9}$$

$$y(t) = x(t) \tag{3.10}$$

根据式（3.7）和式（3.8），确切地离散化得

$$x_{t+1} = e^{-T_s/\tau}x_t + k(1 - e^{-T_s/\tau})u_t \tag{3.11}$$

$$y_t = x_t \tag{3.12}$$

式（3.11）代入式（3.12），然后将采样时间减 1 得

$$y_t = e^{-T_s/\tau}y_{t-1} + k(1 - e^{-T_s/\tau})u_{t-1} \tag{3.13}$$

由一个离散时间传递函数表达差分方程更方便。可以引入一个逆变换算子 z^{-1} 实现，这样有 $z^{-1}y_t = y_{t-1}$、$z^{-2}u_t = u_{t-2}$ 等。使用逆变换算子，离散时间传递函数由式（3.13）所表示的差分方程可以表示为

$$y_t = e^{-T_s/\tau}z^{-1}y_t + k(1 - e^{-T_s/\tau})z^{-1}u_t \tag{3.14}$$

整理这个方程可得

$$G_p(z^{-1}) = \frac{y_t}{u_t} = \frac{k(1 - e^{-T_s/\tau})z^{-1}}{1 - e^{-T_s/\tau}z^{-1}} \tag{3.15}$$

式中，$G_p(z^{-1})$ 表示离散时间传递函数。

该过程的时间延迟可以通过逆变换算子很容易地表示，特别是时间延迟是采样间隔 T_s 的整数倍时。例如，如果一个进程拥有 dT_s 的时间延迟，那么离散时间延迟可以通过一个简单的传递函数表示为 z^{-d}。要判断一个进程是否有时间延迟，建议先写过程中的传递函数 $G_p(z^{-1})$ 然后检查是否 $G_p(z^{-1}=0)=0$。如果是，则该过程有时间延迟。例如，根据式（3.15），$G_p(z^{-1}=0)=0$，所以在这个过程有时间延迟。如果去掉分子的 z^{-1}，则剩下的传递函数是

$$\tilde{G}_p(z^{-1})=\frac{k(1-e^{-T_s/\tau})}{1-e^{-T_s/\tau}z^{-1}}$$

和

$$\tilde{G}_p(z^{-1}=0)=k(1-e^{-T_s/\tau})\neq 0$$

因此，可以说 $\tilde{G}_p(z^{-1})$ 没有时间延迟，所以 $\tilde{G}p(z^{-1})$ 是所谓的传递函数 $G_p(z^{-1})$ 过程中的无延迟部分。$\tilde{G}_p(z^{-1})$ 和 $G_p(z^{-1})$ 的不同是系数 z^{-1}，$G_p(z^{-1})$ 过程有一个采样时间延迟。如果不同的是 z^{-d}，$G_p(z^{-1})$ 过程会有 d 个采样时间延迟。

随着离散时间函数的引入，一阶系统的准确的离散结果可以归纳为以下变换对：

$$\frac{y(s)}{u(s)}=\frac{k}{\tau s+1}\Leftrightarrow\frac{y_t}{u_t}=\frac{k(1-e^{-T_s/\tau})}{1-e^{-T_s/\tau}z^{-1}} \tag{3.16}$$

通过部分分式展开，高阶连续时间传递函数可以写为一阶传递函数的总和。准确的离散然后就可以依据一阶离散结果运行。例如，给定一个二阶传递函数

$$\frac{y(s)}{u(s)}=\frac{k(\tau_0 s+1)}{(\tau_1 s+1)(\tau_2 s+1)}$$

部分分式展开得

$$\frac{y(s)}{u(s)}=\frac{k(\tau_1-\tau_0)}{\tau_1-\tau_2}\,\frac{1}{\tau_1 s+1}+\frac{k(\tau_0-\tau_2)}{\tau_1-\tau_2}\,\frac{1}{\tau_2 s+1} \tag{3.17}$$

首先，式（3.16）应用于式（3.17）右侧的每一项，采样周期为 T_s。然后，将得到的一阶离散传递函数加到一起得出

$$\frac{y_t}{u_t}=\frac{\dfrac{k(\tau_1-\tau_0)}{\tau_1-\tau_2}(1-e^{-T_s/\tau_1})}{1-e^{-T_s/\tau_1}z^{-1}}+\frac{\dfrac{k(\tau_0-\tau_2)}{\tau_1-\tau_2}(1-e^{-T_s/\tau_2})}{1-e^{-T_s/\tau_2}z^{-1}}$$

将其转换为差分方程得出以下结果：

$$y_t+a_1y_{t-1}+a_2y_{t-2}=b_1u_{t-1}+b_2u_{t-2}$$

式中

$$a_1=-e^{-T_s/\tau_1}-e^{-T_s/\tau_2}$$

$$a_2=e^{-T_s/\tau_1}e^{-T_s/\tau_2}$$

$$b_1=k\left(1+\frac{\tau_0-\tau_1}{\tau_1-\tau_2}e^{-T_s/\tau_1}+\frac{\tau_2-\tau_0}{\tau_1-\tau_2}e^{-T_s/\tau_2}\right)$$

$$b_2 = k\left(e^{-T_s(1/\tau_1+1/\tau_2)} + \frac{\tau_0-\tau_1}{\tau_1-\tau_2}e^{-T_s/\tau_2} + \frac{\tau_2-\tau_0}{\tau_1-\tau_2}e^{-T_s/\tau_1}\right)$$

请注意，如果传递函数的分母中有重复的根的话，部分分式展开和随后的离散化将比上面给出的例子更加复杂。

例如 3. 1. 1 考虑一阶微分方程

$$10\frac{dy(t)}{dt} + 2y(t) = 3u(t)$$

式中，$u(t)$ 是分段常数。

用 $T_s=1$ 为这个常微分方程开发一个差分方程和

1）有限差分；

2）准确的离散；

3）通过离散时间传递函数推导出 1 和 2 来表述离散化模型。

1. 通过有限差分解决

$$\frac{dy\ (t)}{dt} \approx \frac{y_t - y_{t-1}}{T_s}$$

因此，原来的差分方程可以写为

$$10\frac{y_t - y_{t-1}}{T_s} + 2y_{t-1} = 3u_{t-1} \tag{3.18}$$

令 $T_s=1$，式（3. 18）得

$$y_t = 0.8y_{t-1} + 0.3u_{t-1} \tag{3.19}$$

2. 通过准确的离散解决

原来的差分方程可以改写为连续的时间传递函数形式：

$$\frac{y(s)}{u(s)} = \frac{3}{10s+2} = \frac{1.5}{5s+1} \tag{3.20}$$

另外，$k=1.5$，$\tau=5$。

把式（3. 13）应用到准确的离散式（3. 20），可以得到

$$y_t = 1.5(1-e^{-0.2T_s})u_{t-1} + e^{-0.2T_s}y_{t-1} \tag{3.21}$$

令 $T_s=1$，式（3. 21）变为

$$y_t = 0.8187y_{t-1} + 0.2719u_{t-1} \tag{3.22}$$

3. 离散时间传递函数

若差分方程源自有限差分法，离散时间传递函数是

$$G_p(z^{-1}) = \frac{0.3z^{-1}}{1-0.8z^{-1}}$$

若差分方程源自准确的离散，离散时间传递函数是

$$G_p(z^{-1}) = \frac{0.2719z^{-1}}{1-0.8187z^{-1}}$$

可以验证的是，它们都有一个样本时间延迟。

3.2　信号

3.2.1　输入信号

线性系统辨识常用的输入信号包括阶跃函数、伪随机或随机二进制序列（RBS），自回归移动平均数序列和正弦曲线的总和。

一个阶跃函数由下式给出：

$$u_t = \begin{cases} 0 & t < 0 \\ u_0 & t \geqslant 0 \end{cases}$$

它通常用于识别实验，如果信噪比（SNR）足够大。阶跃响应包含有关进程动态的信息，例如增益、时间常数和时间延迟，如果噪声可以忽略不计，这些参数可以轻松地根据响应读取。任何一个一阶过程的阶跃响应呈现在图 3.1 中，其中 τ_d 是该过程的时间延迟，τ 为时间常数。

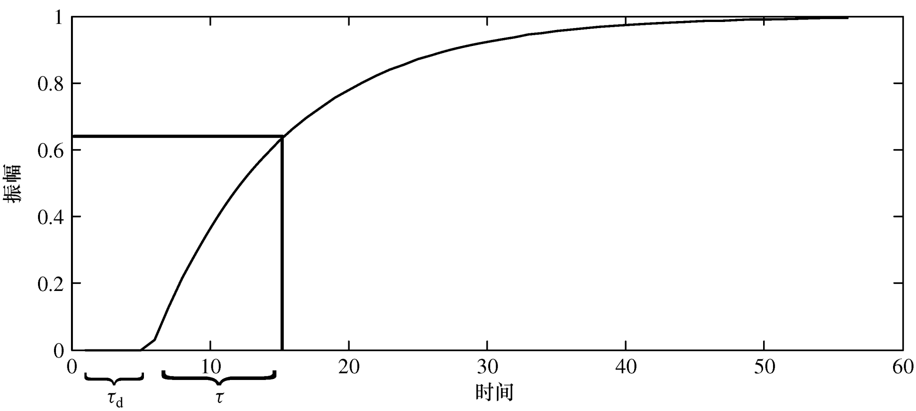


图 3.1　一阶过程的阶跃回应

当在阶跃实验过程中存在显著的噪声时，从阶跃响应提取模型参数比较困难，如图 3.2 所示。在这种情况下输入信号的替代者应被考虑。为了达到了解 RBS 的目的，我们考虑以下情形。

如果由于噪声的响应使得对参数估计的一阶测试不够好，可以考虑重复几个试验步骤，使噪声的影响被平均出来。一个完整阶跃测试通常会持续下去直到响应平息，结果是重复完整的测试步骤会比较冗长且费时。接下来的问题是，它是否需要进行一个完整步骤的测试。我们要考虑一下，通过一步测试我们要估计什么。举一个简单的一阶动态系统为例。一阶模型由 3 个参数来估算：增益、延迟时间和时间常数。估算增益，测试步骤的长度应当至少是相同过程的稳定时间。时间延迟的影响，只能出现在该步测试的初步响应，同时一个完整的阶跃响应不利于更好地确定时间延迟。另一方面，只有当响应达到最终值的 63.2% 时时间

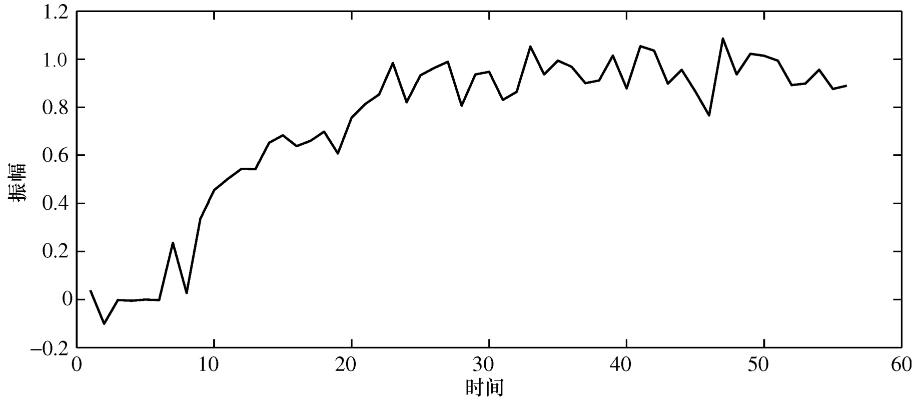


图 3.2　噪声一阶过程的阶跃响应

常数的影响才会显现出来。考虑到估算这些参数需要不同长度的测试步骤，所以有必要进行一个如图 3.3 所示的涵盖 3 个不同时间尺度的实验。考虑到实际的输入设计，该序列不应遵循增益—时间延迟—时间常数，否则它有可能对其中一个比其他的有利。这些序列应该是随机的，以避免实验过程中的任何系统误差，例如，该传感器的性能可能会随着温度的变化而变化，这反过来会随着时间的变化而变化。随机化可以减少系统误差。另外，不同长度的步骤测试应该进行数次，以减少随机噪声的影响，而且该序列必须被完全随机化。这代表了所谓的伪随机或随机二进制信号。

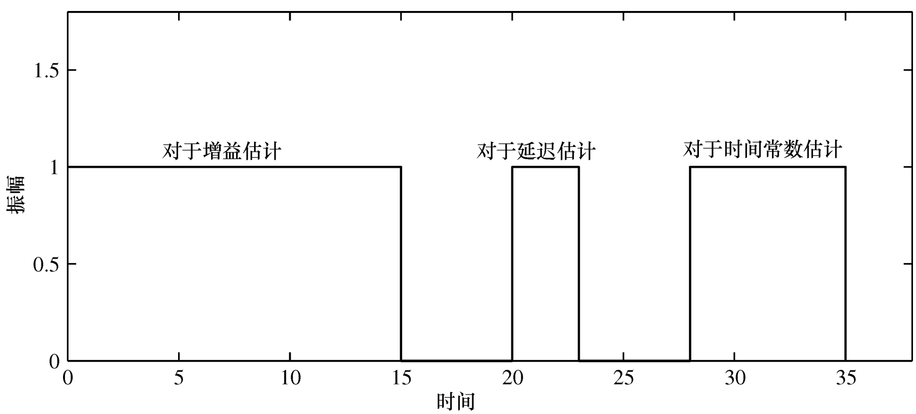


图 3.3　3 个尺度参数的输入设计

正式地说，一个伪随机二进制序列（PRBS）或随机二进制序列（RBS）是一个以某种方式在两者之间变化的信号。伪随机二进制信号是周期性的二进制信号，而随机二进制信号是“非周期”二进制信号，如图 3.4 所示。RBS 是最常用到的对线性系统进行识别实验的输入信号。

时间序列是另一种类型的输入信号。让 e_t 是具有以下属性的白噪声序列：

$$Ee_t = 0$$

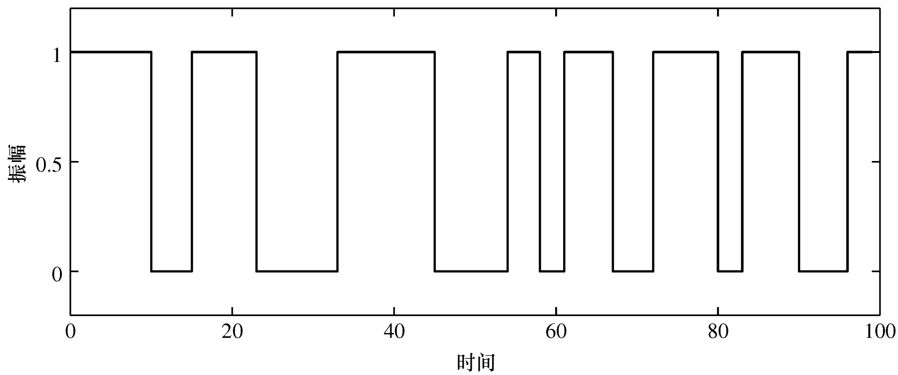


图 3.4 RBS

$$Ee_te_{t-j} = \begin{cases} \lambda^2 & 如果\ j=0 \\ 0 & 如果\ j\neq 0 \end{cases}$$

式中，E 为期望因子。

然后，下面的模型被称为 ARMA（nd，nc）（自回归移动平均模型）随机信号 u_t 的模型，它会生成自回归移动平均序列：

$$u_t + d_1u_{t-1} + \cdots + d_{nd}u_{t-nd} = e_t + c_1e_{t-1} + \cdots + c_{nc}e_{t-nc}$$

式中，nd 和 nc 是 ARMA 模型的阶。

这个模型（差分方程）可以变换为一个离散时间传递函数：

$$\frac{u_t}{e_t} = \frac{1 + c_1z^{-1} + \cdots + c_{nc}z^{-nc}}{1 + d_1z^{-1} + \cdots + d_{nd}z^{-nd}}$$

一个更普遍的 ARIMA（nd，nc）的模型（一个集成的 ARMA）可以写成

$$\frac{u_t}{e_t} = \frac{1 + c_1z^{-1} + \cdots + c_{nc}z^{-nc}}{\Delta(1 + d_1z^{-1} + \cdots + d_{nd}z^{-nd})}$$

其中，$\Delta = (1 - z^{-1})$，操作 $1/\Delta$ 也被称为积分器。时间序列模型考虑由下式给出：

$$\frac{u_t}{e_t} = \frac{1}{\Delta} \tag{3.23}$$

对积分长除得

$$\frac{u_t}{e_t} = \frac{1}{\Delta} = \frac{1}{1 - z^{-1}} = 1 + z^{-1} + z^{-2} + z^{-3} + \cdots$$

因此，式（3.23）可转化为

$$u_t = e_t + e_{t-1} + e_{t-2} + e_{t-3} + \cdots$$

这实际上是对 e_t 积分。式（3.23）也可以替代表示为

$$(1 - z^{-1})u_t = e_t$$

已经给出：

$$u_t = u_{t-1} + e_t$$

也就是说，电流信号 u_t 是以前的值的总和和一个随机的白噪声。这类似于一个喝醉酒的人随机地从当前位置走开。因此，由式（3.23）得到的信号 u_t 也被称为随机游动信号。

ARMA 模型的一种特殊情况是

$$\frac{u_t}{e_t} = \frac{1}{1 + d_1 z^{-1} + \cdots + d_{nd} z^{-nd}}$$

被称为第二阶的自回归模型，记为 AR（nd），然而

$$\frac{u_t}{e_t} = 1 + c_1 z^{-1} + \cdots + c_{nc} z^{-nc}$$

被称为阶 nc 的移动平均模型，记为 MA（nc）（移动平均线）。AR 模型能够直接利用最小二乘法，而另一个被估计时间序列模型涉及更复杂的评估程序。

时间序列序 u_t 可以看作白噪声 e_t 的滤波输出。在一般情况下，时间序列 u_t 可以通过下式获得：

$$u_t = \frac{C(z^{-1})}{D(z^{-1})} e_t$$

式中

$$C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_{nc} q^{-nc}$$
$$D(z^{-1}) = 1 + d_1 z^{-1} + \cdots + d_{nd} q^{-nd}$$

这个过滤 u_t 也称为有色噪声。这是相对于白噪声 e_t 和类似于过滤白色光束产生的彩色光束。对于过滤光的稳定性，对 z^{-1} 而言，$D(z^{-1})$ 的根必须在单位圆外或对 z 而言 $D(z)$ 的根必须在单位圆内。图 3.5 显示了白噪声的轨迹。图 3.6 所示的是过滤白噪声的轨迹，其中过滤由下式给出：

$$u_t = \frac{1 - 0.5 z^{-1}}{1 - 0.98 z^{-1}} e_t$$

可以看出，经过滤的白噪声是比白噪声平滑的。白噪声和过滤白噪声相对于其光谱特性而言已经近似形成数学表达式（将在后面详述），但是它们作为识别信号用于识别实验很难实现。因此，它们的理论价值超过了实际应用。

这里介绍的最后一种输入信号的类型是周期性的输入序列。在混合类正弦信号中，u_t 由下式给出：

$$u_t = \sum_{i=1}^{m} a_i \sin(\omega_i t + \varphi_i)$$

式中，角频率 $\{\omega_j\}$ 是不同的。为了正弦曲线的和，用户必须选择振幅 $\{a_i\}$、频率 $\{\omega_i\}$ 和相位 $\{\varphi_i\}$。图 3.7 所示为这种类型信号的一个例子。

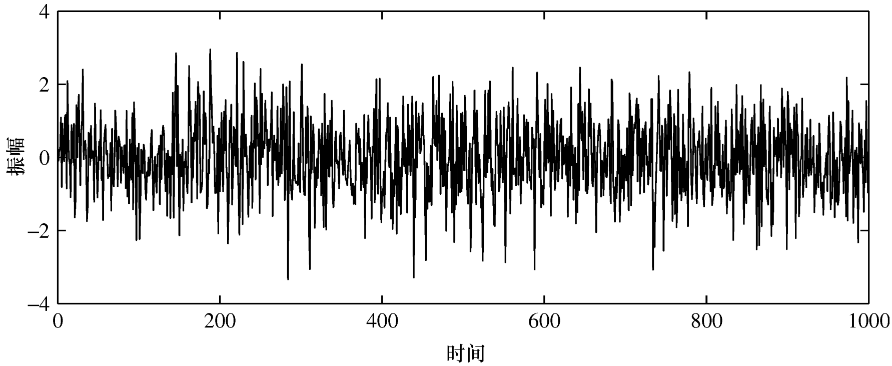


图 3.5 白噪声序列

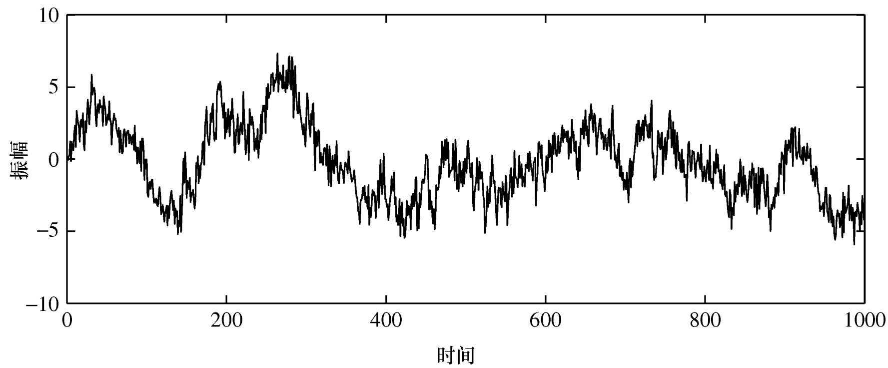


图 3.6 过滤白噪声序列

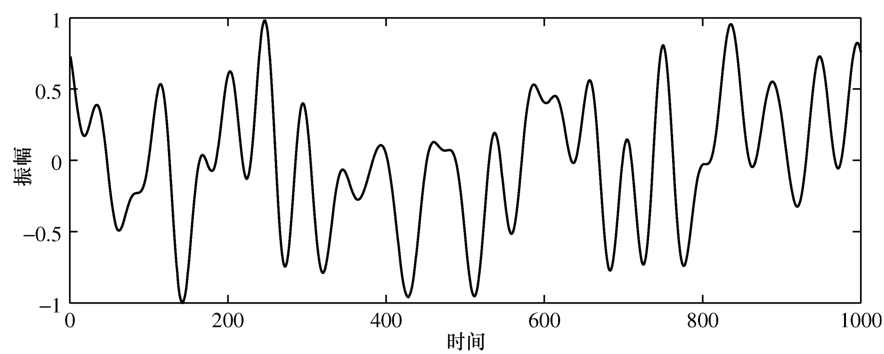


图 3.7 正弦序列总和

3.2.2 信号的光谱特性

对于一个平稳随机过程 u_t，均值 m 和均方差函数 $r_u(\tau)$ 被定义为

$$m = Eu_t$$

$$r_u(\tau) = E(u_{t+\tau} - m)(u_t - m)^{\mathrm{T}}$$

对于一个确定的信号，相应的定义是通过代入 E 和标准化总和极限获得的：

$$m = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} u_t$$

$$r_u(\tau) = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}[u_{t+\tau} - m][u_t - m]^{\tau}$$

对于许多随机过程，这两种定义是等价的。这样的随机过程被称为遍历进程。

一个随机或确定性信号 u_t 的频谱密度 $\varphi_u(\omega)$，可以定义为协方差函数的离散傅里叶变换：

$$\phi_u(\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} r_u(\tau)\mathrm{e}^{-i\tau\omega}$$

$r_u(\tau)$ 也可以通过 $\varphi_u(\omega)$ 的逆变换计算

$$r_u(\tau) = \int_{-\pi}^{\pi}\phi_u(\omega)\mathrm{e}^{i\tau\omega}\mathrm{d}\omega \tag{3.24}$$

输入信号的频谱密度的形状在识别实验的设计中起重要作用。当设计一个输入过程，它对于考虑估计识别一个动作过程的频率响应是很重要的。一个过程的频率响应可以通过博德图来表示。如图 3.8 所示，两个方框是等价的。确定的传递函数模型（上图）等效于识别的博德图（下图）。在原则上，精心设计的输入序列应该使其频谱覆盖所有博德图中的频率范围，过程如图 3.9 所示。不难发现，一个白噪声序列，e_t，在所有频率中具有平坦的频谱。因此，“白”噪声源自于这样一个事实，从白光类推，对于所有的频率它具有相同的幅度。白噪声序列，因此是在理论上用于系统识别的最好的输入信号，但它不是“用户有好的”，因为在实践中难以落实。

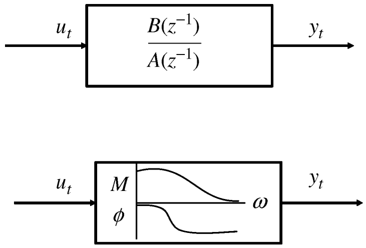


图 3.8　博德图和传递函数的等价性

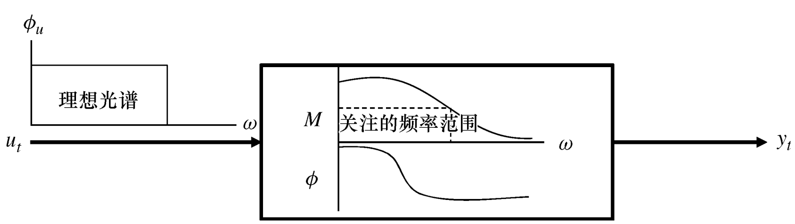


图 3.9　覆盖响应频率的理想输入谱图

当 $\tau=0$、$r_u(0)=\mathrm{Var}(u_t)$ 时，它遵循式（3.24）有

$$\mathrm{Var}(u_t) = \int_{-\pi}^{\pi}\phi_u(\omega)\mathrm{d}\omega$$

这种关系被称为 Parseval 关系。Parseval 关系为时域方差和频域频谱之间提供了一个连接，也就是说，在该地区一个信号的频谱是信号的方差。

对于协方差的伪随机二进制信号 u_t 以及周期 M 和 a 的大小，其协方差函数和谱密度是由 Söderström 和 Stoica（1989）给出的：

$$r_u(\tau)=\begin{cases}a^2 & \tau=0,\pm M,\pm 2M,\cdots\\ -a^2/M & \text{其他}\end{cases}$$

和

$$\phi_u(\omega)=\frac{a^2}{M^2}\left[\delta(\omega)+(M+1)\sum_{k=1}^{M-1}\delta\left(\omega-2\pi\frac{k}{M}\right)\right]$$

当 M 趋近于无穷大时，$\phi_u(\omega)$ 收敛到一个平坦的频谱，类似于白噪声的频谱。

一个 ARMA 过程 u_t 的谱密度由下式给出：

$$\phi_u(\omega)=\frac{\lambda^2}{2\pi}\left|\frac{C(\mathrm{e}^{-i\omega})}{D(\mathrm{e}^{-i\omega})}\right|^2$$

式中，λ^2 是白噪声 e_t 的方差。显然，谱密度的形状取决于 ARMA 模型的参数。ARMA 模型也可以被看做是过滤器。因此，人们可以通过适当形状 u_t 的频谱为一个所需的输入调谐滤波器参数。

例 3.2.1　时间序列模型由下式给出：

$$v_t-0.9v_{t-1}=e_t-0.5e_{t-1}$$

式中，v_t 是时间序列信号；e_t 是驱动白噪声 $\lambda^2=2\pi$ 的方差。推导出信号 v_t 的频谱密度。

差分方程

$$v_t-0.9v_{t-1}=e_t-0.5e_{t-1}\tag{3.25}$$

可改写为

$$v_t=\frac{1-0.5z^{-1}}{1-0.9z^{-1}}e_t\tag{3.26}$$

得到的信号 $v(t)$ 的频谱密度如下：

$$\Phi_v(\omega)=\left|\frac{1-0.5\mathrm{e}^{-\mathrm{j}\omega}}{1-0.9\mathrm{e}^{-\mathrm{j}\omega}}\right|^2\frac{\lambda^2}{2\pi}\tag{3.27}$$

为了得到 $|1-0.5\mathrm{e}^{-\mathrm{j}\omega}|^2$ 和 $|1-0.9\mathrm{e}^{-\mathrm{j}\omega}|^2$，这个复数的下列属性被用到：

$$\begin{aligned}|A+B\mathrm{e}^{\mathrm{j}\omega}|^2&=(A+B\mathrm{e}^{\mathrm{j}\omega})(A+B\mathrm{e}^{-\mathrm{j}\omega})\\&=A^2+B^2+AB(\mathrm{e}^{\mathrm{j}\omega}+\mathrm{e}^{-\mathrm{j}\omega})\\&=A^2+B^2+2AB\cos\omega\end{aligned}\tag{3.28}$$

若 $\lambda^2=2\pi$ 已知，式（3.27）因而可简化为

$$\Phi_v(\omega)=\frac{1.25-\cos\omega}{1.81-1.80\cos\omega}\tag{3.29}$$

3.2.3　输入信号的持续激励

为了说明这个概念，我们以一个简单的例子开始本节。考虑一个二阶与扰动差分方程：

$$y_t = -a_1 y_{t-1} - a_2 y_{t-2} + b_1 u_{t-1} + b_2 u_{t-2} + e_t$$

从 $t=3$ 开始，一组方程可以写为

$$\begin{aligned} y_3 &= -a_1 y_2 - a_2 y_1 + b_1 u_2 + b_2 u_1 + e_3 \\ y_4 &= -a_1 y_3 - a_2 y_2 + b_1 u_3 + b_2 u_2 + e_4 \\ &\vdots \\ y_{N-1} &= -a_1 y_{N-2} - a_2 y_{N-3} + b_1 u_{N-2} + b_2 u_{N-3} + e_{N-1} \\ y_N &= -a_1 y_{N-1} - a_2 y_{N-2} + b_1 u_{N-1} + b_2 u_{N-2} + e_N \end{aligned}$$

将这些等式以矩阵形式给出：

$$\begin{pmatrix} y_3 \\ y_4 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix} = \begin{pmatrix} -y_2 & -y_1 & u_2 & u_1 \\ -y_3 & -y_2 & u_3 & u_2 \\ \vdots & \vdots & \vdots & \vdots \\ -y_{N-2} & -y_{N-3} & u_{N-2} & u_{N-3} \\ -y_{N-1} & -y_{N-2} & u_{N-1} & u_{N-2} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} e_3 \\ e_4 \\ \cdots \\ e_{N-1} \\ e_N \end{pmatrix}$$

以一个紧凑的形式可得

$$\boldsymbol{Y} = \boldsymbol{X\theta} + \boldsymbol{\epsilon}$$

众所周知的最小二乘解得

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$$

$\hat{\boldsymbol{\theta}}$ 存在的唯一的解依赖于（$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$）$^{-1}$ 的存在。如果矩阵 $\boldsymbol{X}$ 的秩小于4，就不可能有解。这可能会发生，例如，当 $\boldsymbol{X}$ 矩阵的列线性相关时因为输入 u_t 设计不良，就有可能发生。即使任何两列之间的相关性是不完全线性的，一个强大的相关性可能仍然膨胀（$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$）$^{-1}$，因此，膨胀变换协方差估计是因为 $\mathrm{Cov}(\hat{\boldsymbol{\theta}}) = \mathrm{Var}(e_t)(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$。在这种情况下，我们说的输入序列不是持续激励。换言之，因为输入信号的设计不周，估计误差可能会无限大。一个可避免（$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$）$^{-1}$ 不良状况的输入信号称为持续激励信号。

通常，一个 n 阶信号 u_t 被说成是持续激励（pe），如果（Söderström 和 Stoica，1989）：

1）存在以下极限：

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_{t+\tau} u_t^T \tag{3.30}$$

和

2）矩阵

$$\boldsymbol{R}_u(n)=\begin{pmatrix} r_u(0) & r_u(1) & \cdots & r_u(n-1) \\ r_u(-1) & r_u(0) & \cdots & r_u(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ r_u(1-n) & \cdots & \cdots & r_u(0) \end{pmatrix} \tag{3.31}$$

是正定的。

对 n 阶线性系统一致估计的一个必要条件是输入信号应该是 $2n$ 阶持续激励的，除了外源性输入（ARX）模型的自回归。要检查这个例外，让我们来看一个简单的 ARX 模型：

$$y_t=ay_{t-1}+bu_{t-1}+e_t$$

式中，u_t 和 e_t 是相互独立的，e_t 是白噪声，方差为 λ^2。

在一般情况下我们可以发现，一阶的持续激励输入信号 u_t 足以实现一个一致的估计 a 和 b。

令

$$\widetilde{\boldsymbol{Y}}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix},\quad \widetilde{\boldsymbol{U}}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix},\quad \boldsymbol{\Phi}=(\widetilde{Y}\ \widetilde{U})$$

$$\boldsymbol{Y}=\begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_{N+1} \end{pmatrix},\quad \boldsymbol{\epsilon}=\begin{pmatrix} e_2 \\ e_3 \\ \vdots \\ e_{N+1} \end{pmatrix},\quad \boldsymbol{\theta}=\begin{pmatrix} a \\ b \end{pmatrix},\quad \hat{\boldsymbol{\theta}}=\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$$

式中，$\boldsymbol{Y}=\boldsymbol{\Phi\theta}+\boldsymbol{\epsilon}$。

如果 $\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$ 是可逆的，那么我们有

$$\begin{aligned}\hat{\boldsymbol{\theta}}&=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi\theta}+\boldsymbol{\epsilon})\\&=\boldsymbol{\theta}+(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}=\boldsymbol{\theta}+\left(\frac{1}{N}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)^{-1}\frac{1}{N}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}\end{aligned}$$

鉴于一阶持续激励，可以证明如下：

1）$\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$ 是可逆的；

2）$\lim\limits_{N\to\infty}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}=0$。

证明：证明 1），从线性代数理论可以知道，$\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$ 的可逆等价于左侧的可逆，当且仅当（i）$\widetilde{U}\neq 0$ 及（ii）对于任意标量的 α，$\widetilde{Y}$ 不等于 $\alpha\widetilde{U}$。可以知道，u_t 是一阶持续激励的，即对于足够大的 N，$\sum\limits_{t=1}^{N}u_t^2\neq 0$。这意味着 $\widetilde{U}\neq 0$。条件（ii）也是如此，因为由于外部白噪声的干扰，$\widetilde{Y}=\alpha\widetilde{U}$ 几乎是不可能的。

另一种解释 1）的真实情况如下：

$$\det\left(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)=\det\begin{pmatrix}\sum_{t=1}^{N}y_t^2 & \sum_{t=1}^{N}y_tu_t\\ \sum_{t=1}^{N}y_tu_t & \sum_{t=1}^{N}u_t^2\end{pmatrix}$$

$$=\sum_{t=1}^{N}y_t^2\sum_{t=1}^{N}u_t^2-\left(\sum_{t=1}^{N}y_tu_t\right)^2$$

众所周知的事实是，当且仅当（$\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$）不等于 0 时，$\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$ 是可逆的。根据柯西（Canchy- Schwarz）不等式：

$$\sum_{t=1}^{N}y_t^2\sum_{t=1}^{N}u_t^2\geqslant\left(\sum_{t=1}^{N}y_tu_t\right)^2$$

并且当且仅当条件（i）和（ii）同时满足等号成立。因此 $\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}$ 是可逆的。

为了证明 2)，我们注意到

$$\lim_{N\to\infty}\left(\frac{1}{N}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}\right)=\lim_{N\to\infty}\begin{pmatrix}\frac{1}{N}\sum_{t=1}^{N}y_te_{t+1}\\ \frac{1}{N}\sum_{t=1}^{N}u_te_{t+1}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

最后一个等式成立，因为 u_t 和 e_t 被称为是相互独立的，y_t 可以不依赖于未来的白噪声 e_{t+1}，因此

$$\lim_{N\to\infty}\hat{\boldsymbol{\theta}}=\lim_{N\to\infty}\left(\boldsymbol{\theta}+\left(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}\right)$$

$$=\boldsymbol{\theta}+\lim_{N\to\infty}\left(\frac{1}{N}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)^{-1}\lim_{N\to\infty}\left(\frac{1}{\boldsymbol{N}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\epsilon}\right)=\boldsymbol{\theta}$$

式中

$$\lim_{N\to\infty}\left(\frac{1}{N}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)=\lim_{N\to\infty}\begin{pmatrix}\frac{1}{N}\sum_{t=1}^{N}y_t^2 & \frac{1}{N}\sum_{t=1}^{N}y_tu_t\\ \frac{1}{N}\sum_{t=1}^{N}y_tu_t & \frac{1}{N}\sum_{t=1}^{N}u_t^2\end{pmatrix}$$

$$=\begin{pmatrix}r_y(0) & r_{yu}(0)\\ r_{uy}(0) & r_u(0)\end{pmatrix}$$

已被证明是可逆的。

例 3.2.2　考虑一个设计输入实验过滤白噪声通过下面的滤波：

$$u_t=(1+cz^{-1})e_t$$

其中，e_t 为白噪声，方差为 λ^2。讨论这个输入 u_t 是什么样的持续激励顺序。

改写

$$u_t=(1+cz^{-1})e_t=e_t+ce_{t-1}\tag{3.32}$$

由于 e_t 是白噪声，下列性质成立：

$$E[u_t]=0$$

$$E[e_{t+i}e_{t+i}^T]=\lambda^2$$

$$E[e_{t+i}e_{t+j}^T]=0 \quad i\neq j$$

如果式（3.30）和式（3.31）都满足信号 u_t 是持续激励的 n 阶微分方程。为此，我们要评估的协方差函数 $r_u(\tau)$ 为 τ 的不同的值：

$$\begin{aligned} r_u(0) &= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}u_t u_t^T \\ &= E[u_t u_t^T] \\ &= E[(e_t+ce_{t-1})(e_t+ce_{t-1})^T] \\ &= E[e_t e_t^T]+2cE[e_t e_{t-1}^T]+c^2E[e_{t-1}e_{t-1}^T] \\ &= \lambda^2(1+c^2) \end{aligned} \tag{3.33}$$

$$\begin{aligned} r_u(1)=r_u(-1) &= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}u_{t+1}u_t^T \\ &= E[u_{t+1}u_t^T] \\ &= E[(e_{t+1}+ce_t)(e_t+ce_{t-1})^T] \\ &= E[e_{t+1}e_t^T]+cE[e_{t+1}e_{t-1}^T]+cE[e_t e_t^T]+c^2E[e_t e_{t-1}^T] \\ &= \lambda^2 c \end{aligned} \tag{3.34}$$

另外，对于 τ 大于等于 2：

$$\begin{aligned} r_u(\tau)=r_u(-\tau) &= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}u_{t+\tau}u_t^T \\ &= E[u_{t+\tau}u_t^T] \\ &= E[(e_{t+\tau}+ce_{t+\tau-1})(e_t+ce_{t-1})^T] \\ &= E[e_{t+\tau}e_t^T]+cE[e_{t+\tau}e_{t-1}^T] \\ &\quad +cE[e_{t+\tau-1}e_t^T]+c^2E[e_{t+\tau-1}e_{t-1}^T] \\ &= 0 \end{aligned} \tag{3.35}$$

因此，$\boldsymbol{R}_u(n)$ 的矩阵，可以构造如下：

$$\boldsymbol{R}_u(n)=\lambda^2\begin{pmatrix} 1+c^2 & c & & & \\ c & 1+c^2 & c & & \\ & c & \ddots & \ddots & \\ & & \ddots & \ddots & c \\ & & & c & 1+c^2 \end{pmatrix} \tag{3.36}$$

显然，$\boldsymbol{R}_u(n)$ 是一个对称三角矩阵。其中，描述三角矩阵的正定性的最基本的方法是基于矩阵的正定性，被称为希尔维斯特的绕圈（Sylvester's Criterium）。

实对称三角矩阵 $\boldsymbol{R}_u(n)$ 只有当最主要的小数 $\boldsymbol{R}_u(k)$ 对于 $k=1, \cdots, n$ 是正定的：

$$\det\boldsymbol{R}_u(1)=\lambda^2(1+c^2)>0 \tag{3.37}$$

$$\begin{aligned}\det\boldsymbol{R}_u(2)&=\lambda^4\begin{vmatrix}1+c^2 & c\\ c & 1+c^2\end{vmatrix}\\ &=\lambda^4((1+c^2)^2-c^2)\\ &=\lambda^4(1+c^2+c^4)>0\end{aligned} \tag{3.38}$$

$$\begin{aligned}\det\boldsymbol{R}_u(3)&=\lambda^6\begin{vmatrix}1+c^2 & c & 0\\ c & 1+c^2 & c\\ 0 & c & 1+c^2\end{vmatrix}\\ &=\lambda^6((1+c^2)(1+c^2+c^4)-c^2(1+c^2))\\ &=\lambda^6(1+c^2+c^4+c^6)>0\end{aligned} \tag{3.39}$$

此外，它可以证明

$$\det\boldsymbol{R}_u(n)=\lambda^{2n}(1+c^2+c^4+\cdots+c^{2n})>0 \tag{3.40}$$

因此，可以得出结论，该信号 u_t 对于任意阶是持续激励的。

3.2.4　输入设计

理论上理想的输入是白噪声，因为在所有的频率它都能够激发，这相当于在所有阶都是持续激励的。然而，实际过程中宽度总是有限的，没必要激发所有的频率。实际上，输入信号有太宽的频率范围不仅会过度损耗而且会激发一些不必要的动因。因此，在动态实验中设计恰当的输入信号非常关键。考虑一阶过程：

$$G(s)=\frac{K}{\tau s+1}$$

它的宽度可以确定为 $1/\tau$。所以设计输入信号的宽度应该大于 $1/\tau$，表示为 k/τ，其中 k 大于 1。MATLAB™的系统识别工具箱有一个功能称为 id 输入可以用来在给定的宽度设计一个输入信号。要使用此功能，必须通过奈奎斯特频率对频率归一化，它是采样频率的一半，即 π/T_s。这样，输入信号的宽度对频率归一化后表示应为

$$\omega_{\max}=\frac{kT_s}{\pi\tau}$$

一般情况下，一个过程中往往有有益的低频动态变化，因此对于一阶动态输入信号的频率范围应为 $0\leqslant\omega\leqslant kT_s/\pi\tau$。作为一个经验法则，它通常要求 k 取 2 ~3 的任意值。

3.3　模型

3.3.1　线性模型

动态过程模型可被大致分为线性和非线性模型。线性模型可以进一步分为参数

和非参数化模型。

一般参数模型由下式给出 Ljung（1999）：

$$y_t = G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t \tag{3.41}$$

式中，$G_p(z^{-1};\theta)$ 是过程模型；$G_l(z^{-1};\theta)$ 是扰动模型。假设以下内容：

- $G_l^{-1}(z^{-1};\theta)$ 和 $G_l^{-1}(z^{-1};\theta)\ G_p(z^{-1};\theta)$ 是渐进稳定的；
- $G_p(0;\theta)=0$，$G_l(0;\theta)=I$

然而，$G_l(z^{-1};\theta)$ 在实践中不仅是渐进稳定的。正如下面将要讨论的，临界稳定的 $G_l(z^{-1};\theta)$ 可用于描述数据中的扰动漂移。某些广泛使用的机型都在下面讨论。

此类最普遍的模式由式（3.41）所示，模型是预测误差（PE）的模型，通过下式描述：

$$A(z^{-1})y_t = \frac{B(z^{-1})}{F(z^{-1})}u_t + \frac{C(z^{-1})}{D(z^{-1})}e_t$$

式中

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_{na} z^{-na}$$
$$B(z^{-1}) = b_1 z^{-1} + \cdots + b_{nb} z^{-nb}$$
$$C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_{nc} z^{-nc}$$
$$D(z^{-1}) = 1 + d_1 z^{-1} + \cdots + d_{nd} z^{-nd}$$
$$F(z^{-1}) = 1 + f_1 z^{-1} + \cdots + f_{nf} z^{-nf}$$

要估计的参数是

$$\theta = [a_1, \cdots, a_{na}, b_1, \cdots, b_{nb}, c_1, \cdots, c_{nc}, d_1, \cdots, d_{nd}, f_1, \cdots, f_{nf}]^{\mathrm{T}}$$

该模型可以进一步写为

$$y_t = \frac{B(z^{-1})}{A(z^{-1})F(z^{-1})}u_t + \frac{C(z^{-1})}{A(z^{-1})D(z^{-1})}e_t$$

一个 PE 模型可以分别在传递和干扰函数中灵活指定任何多项式的分子和分母，此外对于在分母中一个共同的因子 $A(z^{-1})$，具有传递和干扰传输功能。

通过让 $C(z^{-1})=D(z^{-1})=F(z^{-1})=1$ 在 PE 模型中，得到一个 ARX 模型为

$$A(z^{-1})y_t = B(z^{-1})u_t + e_t$$

以传递函数的形式，一个 ARX 模型可以写成

$$y_t = \frac{B(z^{-1})}{A(z^{-1})}u_t + \frac{1}{A(z^{-1})}e_t$$

因此，ARX 模型的结构受限制，低阶 ARX 模型动态过程难以捕捉。然而，高阶 ARX 模型可用于近似大多数其他模型结构。事实上，ARX 模型可以估计出，通过紧密型最小二乘解使一个 ARX 结构有吸引力地实践，除了它的高阶。它经常被用来快速或初始解决模型识别问题。

例如，考虑一个 Box-Jenkins（BJ）模型（稍后给出）所描述的系统：

$$y_t=\frac{B(z^{-1})}{F(z^{-1})}u_t+\frac{C(z^{-1})}{D(z^{-1})}e_t \tag{3.42}$$

重写模型：

$$\begin{aligned}F(z^{-1})D(z^{-1})y_t &= B(z^{-1})D(z^{-1})u_t+C(z^{-1})F(z^{-1})e_t\\ &= B(z^{-1})D(z^{-1})u_t+\frac{1}{\dfrac{1}{C(z^{-1})F(z^{-1})}}e_t\end{aligned} \tag{3.43}$$

如果$\frac{1}{C(z^{-1})F(z^{-1})}$是稳定的，也就是说，$C(z^{-1})F(z^{-1})$ 的根是在单位圆外，它可以展开如下：

$$\frac{1}{C(z^{-1})F(z^{-1})}=g_0+g_1z^{-1}+g_1z^{-2}+\cdots=\sum_{i=0}^{\infty}g_iz^{-i} \tag{3.44}$$

g_m 足够接近为 0 时，可以截断求和：

$$\frac{1}{C(z^{-1})F(z^{-1})}\approx\sum_{i=0}^{m}g_iz^{-i} \tag{3.45}$$

定义 $E(z^{-1})=\sum_{i=0}^{m}g_iz^{-i}$，式（3.43）的 BJ 模型可以近似为

$$F(z^{-1})D(z^{-1})E(z^{-1})y_t=B(z^{-1})D(z^{-1})E(z^{-1})u_t+e_t \tag{3.46}$$

这是一个 ARX 模型。

ARX 扰动模型的分子行松弛，也就是说，用多项式 $C(z^{-1})$ 替换 1，产生自回归移动平均线与外源性输入（ARMAX）模型。一个 ARMAX 模型由下式给出：

$$y_t=\frac{B(z^{-1})}{A(z^{-1})}u_t+\frac{C(z^{-1})}{A(z^{-1})}e_t$$

ARMAX 模型的差分方程形式为

$$A(z^{-1})y_t=B(z^{-1})u_t+C(z^{-1})e_t$$

因此，一个 ARMAX 模型强制过程和干扰模型有一个共同的特征，但其分子是灵活的。一个 ARMAX 模型相当于一个状态空间模型，便于卡尔曼滤波器的设计。在 ARMAX 模型中稍微更一般的形式是集成 ARMAX 模型，即 ARIMAX。

一个 ARIMAX 模型由下式给出：

$$A(z^{-1})y_t=B(z^{-1})u_{t-1}+\frac{C(z^{-1})}{\Delta}e_t \tag{3.47}$$

式中

$$\Delta=1-z^{-1}$$

这个模型通常用于预测控制的设计中。卡尔曼预测可以直接来自于所识别的 ARIMAX 模型，如以下内容所述。

该 ARIMAX 模型中相同的阶数（n）的所有多项式可写为

$$\begin{aligned}&(1-z^{-1})(y_t+a_1y_{t-1}+\cdots+a_ny_{t-n})\\ &=b_1\Delta u_{t-1}+\cdots+b_n\Delta u_{t-n}+e_t+c_1e_{t-1}+\cdots+c_ne_{t-n}\end{aligned}$$

或

$$\begin{aligned} & y_t + a'_1 y_{t-1} + \cdots + a'_n y_{t-n} + a'_{n+1} y_{t-n-1} \\ & = b_1 \Delta u_{t-1} + \cdots + b_n \Delta u_{t-n} + e_t + c_1 e_{t-1} + \cdots + c_n e_{t-n} \end{aligned}$$

式中

$$\begin{aligned} a'_1 &= a_1 - 1 \\ a'_2 &= a_2 - a_1 \\ &\vdots \\ a'_n &= a_n - a_{n-1} \\ a'_{n+1} &= -a_n \end{aligned}$$

状态空间创新形式的实现可以写成（Astrom 和 Wittenmark，1990）

$$x_{t+1} = \boldsymbol{A} x_t + \boldsymbol{B} \Delta u_t + \boldsymbol{L} e_t \tag{3.48}$$

$$y_t = \boldsymbol{C} x_t + e_t \tag{3.49}$$

式中

$$\boldsymbol{A} = \begin{pmatrix} -a'_1 & 1 & 0 & \cdots & 0 \\ -a'_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ -a'_n & 0 & 0 & \cdots & 1 \\ -a'_{n+1} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\boldsymbol{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ 0 \end{pmatrix} \qquad \boldsymbol{L} = \begin{pmatrix} c_1 - a'_1 \\ c_2 - a'_2 \\ \vdots \\ c_n - a'_n \\ -a'_{n+1} \end{pmatrix}$$

$$\boldsymbol{C} = (1 \quad 0 \quad \cdots \quad 0)$$

式中，$\boldsymbol{L}$ 是被称为卡尔曼滤波增益。

从式（3.49）求解 e_t，把它带入式（3.48）中，结果是卡尔曼预测为

$$x_{t+1} = \boldsymbol{A} x_t + \boldsymbol{B} \Delta u_t + \boldsymbol{L}(y_t - \boldsymbol{C} x_t)$$

式中，x_t 被视为预测的状态。

过程和干扰模型在 ARMAX 模型中分母是相同的。删除此约束产生一个更灵活的模型结构，称为 BJ 模型：

$$y_t = \frac{B(z^{-1})}{F(z^{-1})} u_t + \frac{C(z^{-1})}{D(z^{-1})} e_t$$

这是在实践中一个相当普遍的模型，其中从其过程来看，扰动可以有一个完全不同的模型。BJ 模型结构是最常见的，如果同时有兴趣使用过程和干扰模型。

PE 模式的一个特殊情况是输出误差（OE）模型，它由下式给出：

$$y_t = \frac{B(z^{-1})}{F(z^{-1})}u_t + e_t$$

虽然 OE 模型使用 e_t 或白噪声作为干扰，在其他干扰的存在的情况下，OE 模型结构可以用来估计过程模型。因为在 OE 模型中被估计的参数是相对较少的，如果对估计的扰动模型不感兴趣的话，在实践中它通常是一个好的模型结构的选择。

线性 l 输入和 m 输出或多重输入和多输出（MIMO）模式，可被分离成 m 个多输入单输出（MISO）模型：

$$A_i(z^{-1})y_{i,t} = \frac{B_{i1}(z^{-1})}{F_{i1}(z^{-1})}u_{1,t} + \cdots + \frac{B_{il}(z^{-1})}{F_{il}(z^{-1})}u_{l,t} + \frac{C_i(z^{-1})}{D_i(z^{-1})}e_{i,t} \quad i = 1, \cdots, m$$

按照相同的程序，作为单输入单输出（SISO）模式的结构，可以构造 MISO ARX、ARMAX MISO、MISO BJ 和 MISO OE。

对于 MIMO 模型，最方便的模型结构是状态空间模型结构。对于系统辨识，状态空间模型通常通过创新或卡尔曼预测的形式确定：

$$x_{t+1} = \boldsymbol{A}(\theta)x_t + \boldsymbol{B}(\theta)u_t + \boldsymbol{K}(\theta)e_t \tag{3.50}$$

$$y_t = \boldsymbol{C}(\theta)x_t + e_t \tag{3.51}$$

式中，$\boldsymbol{A}$、$\boldsymbol{B}$ 和 $\boldsymbol{C}$ 是系统矩阵；e_t 在这里也被称为创新顺序；$\boldsymbol{K}$ 是卡尔曼预测增益。由于零阶保持采样，二维矩阵通常省略（因此，引入一个采样延时）。

状态空间模型，即式（3.50）和式（3.51），给出也可以转入的传递函数模型：

$$y_t = G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t$$

式中

$$G_p(z^{-1};\theta) = C(\theta)[zI - A(\theta)]^{-1}B(\theta)$$

$$G_l(z^{-1};\theta) = I + C(\theta)[zI - A(\theta)]^{-1}K(\theta)$$

可以看到，$G_p(z^{-1},\ \theta)$ 和 $G_l(z^{-1},\theta)$ 具有相同的分母，表明这是一个 ARMAX 模型。因此，该状态模型相当于一个 ARMAX 模型。

因此，状态空间模型和传递函数模型（或差分方程模型）彼此之间很容易变换。

3.3.2　非线性模型

非线性动态模型通常是由非线性状态空间模型来表示的，如果模型是从第一原理推导：

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x(t), u(t))$$

$$y = h(x(t))$$

有限差分法常被用来分离连续时间状态空间模型，使其变为离散时间状态空间模型，例如

$$x_{t+1} = x_t + T_s f(x_t, u_t)$$
$$y_t = h(x_t)$$

其他离散的方法包括，例如，正交配置法、Galerikin 方法和以平整度为基础的方法。其中一些将在下面的内容进一步讨论。

系统识别流行的输入输出的非线性模型都是非线性 ARMAX（NARMAX）系列：

$$y_t = f(y_{t-1}, \cdots, y_{t-p}, u_t, \cdots, e_{t-q}, e_{t-1}, \cdots, e_{t-r}) + e_t$$

这些模型构成了广泛的一类，包括特殊情况下很多其他类非线性离散时间的模型，诸如非线性 ARX 模型（NARX）。NARX 模型由以下结构给出：

$$y_t = f(y_{t-1}, \cdots, y_{t-p}, u_t, \cdots, u_{t-q}) + e_t$$

非线性 AARX 模型与外部输入被定义为

$$y_t = \sum_{i=1}^{p} \alpha_i f_i(y_{t-i}) + \sum_{j=0}^{q} \beta_j g_j(u_{t-j}) + e_t$$

式中，函数 $\{f_i\ (.)\}$ 和 $\{g_i\ (.)\}$ 是非线性函数。

该 NAARX 模型可以使用回归分析方法来估计，因为它是一个线性的参数模型。

一类特殊 NAARX 模型广泛使用 Hammerstein 模型：

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \sum_{j=1}^{q} b_j f(u_{k-j}) + e_t$$

Hammerstein 模型常被用来表示静态非线性输入，如非线性执行器。

更普遍的 Hammerstein 模型被称为面向块的 Hammerstein 模型：

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \sum_{l=0}^{r} \sum_{j=1}^{qt} b_{l,j} f_l(u_{k-j}) + e_t$$

另一种常用的静态非线性模型是 Wiener 模型：

$$z_t = \sum_{i=1}^{p} a_i z_{t-i} + \sum_{j=0}^{q} b_j u_{k-j} + e_t$$
$$y_t = f(z_t)$$

即输出是非线性的（例如，传感器的非线性）。

线性脉冲响应模型对应是的 Volterra 模型：

$$y_t = y_0 + \sum_{j=0}^{\infty} a_j u_{t-j} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{i,j} u_{t-i} u_{t-j} + \sum_{l=0}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} c_{l,i,j} u_{t-l} u_{t-i} u_{t-j} + \cdots$$

实际应用中需要，这一系列被截断为数量有限的项，这些截断级数被称为有限的 Volterra 模型。一个衰退的记忆系统可以用有限的 Volterra 模型来很好地近似。“衰退记忆的条件”意味着“遥远的过去对现在的影响可以忽略不计”。

3.4　注解和参考文献

用于将连续时间系统变换为离散时间表示的主要参考是 Kwakernaak 和 Sivan（1972）、Seborg 等人（1989）以及 Huang 和 Kadali（2008）（施普林格出版社）。

输入信号和时间序列的主要参考是 Söderström 和 Stoica（1989）。时间序列分析的经典理论可以在 Box 和 Jenkins（1976）的研究中找到。

对于线性动态模型结构的引用是 Huang 和 Kadali（2008）（施普林格出版社）、Söderström 和 Stoica（1989）以及 Ljung 的（1999）的研究。非线性动态模型结构最初被 Pearson 和 Ogunnaike（1997）讨论。

本章介绍的例子基于第一作者在阿尔伯塔大学提供的系统识别过程 CHE662 的讲义，有的实例的解决方案来源于 Khatibisepehr 等人的研究（2009-2011）。

第 4 章　系统辨识 Ⅱ

4.1　回归分析

系统辨识是一个从实验数据出发对动态系统进行建模的领域，通常使用如第 3 章介绍的离散时间模型或差分方程。然而，在系统辨识中，离散时间模型不是从原始的微分方程变换而来。相反，该离散时间模型是直接通过实验数据建立的。通常是通过最小二乘法进行线性参数模型的参数估计，而对于非线性估计的问题，采用数值优化法。我们通过举例说明如何用最小二乘法估计一个简单动态模型的参数来展开本章。

4.1.1　利用外生输入模型的移动平均自动回归

考虑受噪声影响的差分方程：

$$y_t = -a_1 y_{t-1} - a_2 y_{t-2} - \cdots - a_{na} y_{t-na} + b_1 u_{t-1} + b_2 u_{t-2} + \cdots + b_{nb} u_{t-nb} + e_t \tag{4.1}$$

回想一下，白噪声具有如下性质：

$$E(e_t) = 0$$

$$Ee_t e_{t-j} = \begin{cases} \lambda^2 & j=0 \\ 0 & j \neq 0 \end{cases}$$

结果中，e_t表示除额外说明白噪声具有上述性质。本书中，我们规定多项式的阶数：一个多项式 $A(z^{-1})$ 的阶数记为 na 和 $B(z^{-1})$ 的阶数记为 nb，依此类推。例如，式（4.1）可以整理为

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_{na} y_{t-na} = b_1 u_{t-1} + b_2 u_{t-2} + \cdots + b_{nb} u_{t-nb} + e_t$$

方程可以进一步简化为

$$A(z^{-1}) y_t = B(z^{-1}) u_t + e_t$$

式中

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_{na} z^{-na}$$

$$B(z^{-1}) = b_1 z^{-1} + \cdots + b_{nb} z^{-nb}$$

这就是第 3 章所介绍的 ARX 模型。

系统辨识的任务就是用植入的输入和输出数据 $\{u_1, u_2, \cdots, u_N\}$ 和 $\{y_1, y_2, \cdots, y_N\}$ 估计参数 $\{a_1, \cdots, a_{na}\}$ 和 $\{b_1, \cdots, b_{nb}\}$，其中 N 表示采样数据的数目。

再考虑一个二阶 ARX 模型：

$$y_t = -a_1 y_{t-1} - a_2 y_{t-2} + b_1 u_{t-1} + b_2 u_{t-2} + e_t \tag{4.2}$$

简化，记为

$$\boldsymbol{\varphi}(t)=\begin{pmatrix}-y_{t-1}\\-y_{t-2}\\u_{t-1}\\u_{t-2}\end{pmatrix}\quad \boldsymbol{\theta}=\begin{pmatrix}a_1\\a_2\\b_1\\b_2\end{pmatrix}$$

式（4.2）可写作

$$\boldsymbol{y}_t=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+\boldsymbol{e}_t \tag{4.3}$$

假设经输入数列 $\{u_1, u_2, \cdots, u_{N-1}, u_N\}$ 激发后的系统，则处理响应数据 $\{y_1, y_2, \cdots, y_{N-1}, y_N\}$ 可以求出。式（4.3）从 $n=3$ 开始，可以写作：

$$\begin{aligned}y_3&=\varphi^{\mathrm{T}}(3)\theta+e_3\\y_4&=\varphi^{\mathrm{T}}(4)\theta+e_4\\&\vdots\\y_{N-1}&=\varphi^{\mathrm{T}}(N-1)\theta+e_{N-1}\\y_N&=\varphi^{\mathrm{T}}(N)\theta+e_N\end{aligned}$$

以矩阵形式记录代数方程，得到

$$\begin{pmatrix}y_3\\y_4\\\vdots\\y_{N-1}\\y_N\end{pmatrix}=\begin{pmatrix}\varphi^{\mathrm{T}}(3)\\\varphi^{\mathrm{T}}(4)\\\vdots\\\varphi^{\mathrm{T}}(N-1)\\\varphi^{\mathrm{T}}(N)\end{pmatrix}\begin{pmatrix}a_1\\a_2\\b_1\\b_2\end{pmatrix}+\begin{pmatrix}e_3\\e_4\\\vdots\\e_{N-1}\\e_N\end{pmatrix}$$

可以写成紧凑的形式：

$$\boldsymbol{Y}=\boldsymbol{\Phi}\boldsymbol{\theta}+\boldsymbol{\epsilon}$$

$\boldsymbol{\theta}$ 的估计可以通过线性回归方法得出。

4.1.2　线性回归

考虑一个带参数 $\boldsymbol{\theta}$ 的一般线性回归模型：

$$\boldsymbol{y}_t=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+\boldsymbol{e}_t \tag{4.4}$$

式中，$\boldsymbol{y}_t$ 是被测量；$\boldsymbol{\varphi}(t)$ 为已知大小的矢量；$\boldsymbol{\theta}$ 是矢量参数；$\boldsymbol{e}_t$ 是白噪声。

线性回归问题是从测量 y_1，$\varphi(1)$，…，y_N，$\varphi(N)$ 中找到参数矢量 $\boldsymbol{\theta}$ 的估计值 $\hat{\boldsymbol{\theta}}$。给出这些分量，得到线性方程系统为

$$\begin{aligned}y_1&=\varphi^{\mathrm{T}}(1)\theta+e_1\\y_2&=\varphi^{\mathrm{T}}(2)\theta+e_2\\&\vdots\\y_N&=\varphi^{\mathrm{T}}(N)\theta+e_N\end{aligned}$$

这可以写成矩阵形式为

$$\boldsymbol{Y}=\boldsymbol{\Phi}\boldsymbol{\theta}+\boldsymbol{\epsilon}$$

式中

$$\boldsymbol{Y}=\begin{pmatrix}y_1\\ \vdots\\ y_N\end{pmatrix}\quad \boldsymbol{\Phi}=\begin{pmatrix}\varphi^{\mathrm{T}}(1)\\ \vdots\\ \varphi^{\mathrm{T}}(\mathrm{N})\end{pmatrix}\quad \boldsymbol{\epsilon}=\begin{pmatrix}e_1\\ \vdots\\ e_N\end{pmatrix}$$

并且

$$E\ [\boldsymbol{\epsilon}]=0\quad \mathrm{Cov}\ [\boldsymbol{\epsilon}]=E\ [\boldsymbol{\epsilon\epsilon}^{\mathrm{T}}]=\lambda^2\boldsymbol{I}$$

最小二乘估计的 $\boldsymbol{\theta}$ 被定义为矢量 $\hat{\boldsymbol{\theta}}$，最大限度地减少损失函数。损失函数可写为

$$\boldsymbol{J}(\theta)=\frac{1}{N}(\boldsymbol{Y}-\boldsymbol{\Phi\theta})^{\mathrm{T}}(\boldsymbol{Y}-\boldsymbol{\Phi\theta})\tag{4.5}$$

关于 $\boldsymbol{\theta}$ 求导，然后使导数等于零，解由下式给出：

$$\hat{\boldsymbol{\theta}}=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}\tag{4.6}$$

$V(\theta)$ 的相应的最小值是

$$\min_{\theta}\boldsymbol{J}(\theta)=\boldsymbol{J}(\hat{\theta})=\frac{1}{N}[\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}-\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}]$$

最小二乘估计的 $\hat{\boldsymbol{\theta}}$ 也可写作

$$\hat{\boldsymbol{\theta}}=\left[\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right]^{-1}\left[\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{y}_t\right]$$

4.1.3　线性回归分析

使数据满足给定的一个真实模型：

$$y_t=\varphi^{\mathrm{T}}(t)\theta_0+e_t$$

基于假设模型的回归方程为

$$y_t=\varphi^{\mathrm{T}}(t)\theta+e_t$$

其矩阵形式为

$$\boldsymbol{Y}=\boldsymbol{\Phi\theta}+\boldsymbol{\epsilon}$$

最小二乘估计的 $\hat{\boldsymbol{\theta}}$ 通过式（4.6）计算。残差定义为实际输出与预计输出之差：

$$\boldsymbol{\varepsilon}=\boldsymbol{Y}-\hat{\boldsymbol{Y}}=\boldsymbol{Y}-\boldsymbol{\Phi}\hat{\boldsymbol{\theta}}$$

计算预测的均方误差（MSE）为

$$\mathrm{MSE}=\frac{\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon}}{N-p}\tag{4.7}$$

式中，p 为估计参数或 $\hat{\theta}$ 的维数。

然后，满足下面的性质：

1）$\hat{\boldsymbol{\theta}}$ 为 θ_0 的无偏估计。

2）$\hat{\boldsymbol{\theta}}$ 的协方差矩阵由下式给出：

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}})=\lambda^2(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}$$

λ^2 的无偏估计是由 MSE 通过式（4.7）计算给出的，将在后面讨论这些性质。已知 Cov（$\hat{\boldsymbol{\theta}}$）对于计算估计的标准差是必要的，以及参数的置信区间。每个估计参数的标准差是由 Cov（$\hat{\boldsymbol{\theta}}$）的对角元素的二次方根给出的。即 $\boldsymbol{C}=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}$。利用 MATLAB®，$\hat{\boldsymbol{\theta}}$ 的标准差由下式给出：

$$\mathrm{SE}(\hat{\boldsymbol{\theta}})=\lambda\sqrt{\mathrm{diag}(\boldsymbol{C})}$$

由于 λ^2 通常是未知的，它可以通过 MSE 来近似。每个独立元素 $\boldsymbol{\theta}$ 上的（$1-\alpha$)% 置信区间可以由下式确定：

$$\hat{\boldsymbol{\theta}}-t_{\alpha/2,N-p}\mathrm{SE}(\hat{\boldsymbol{\theta}})\leqslant\boldsymbol{\theta}\leqslant\hat{\boldsymbol{\theta}}+t_{\alpha/2,N-p}\mathrm{SE}(\hat{\boldsymbol{\theta}})$$

式中，$t_{\alpha/2,N-p}$ 可以从统计教材中的 t 表中查到；α 是用户确定的显著性水平（通常 $\alpha=0.05$ 或 $\alpha=0.01$）。

4.1.4　加权最小二乘法

到目前为止，已假设干扰 e_t 为白噪声。现在，考虑当它为有色噪声的情况。也就是说在回归方程中 e_t 用有色噪声 v_t 替换，定义如下：

$$\boldsymbol{v}=\begin{bmatrix}v_1\\ \vdots\\ v_N\end{bmatrix}$$

式中

$$E[\boldsymbol{v}]=0\quad E[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}]=\boldsymbol{R}\neq\lambda^2\boldsymbol{I}$$

式中，协方差矩阵 $\boldsymbol{R}$ 是正定矩阵；$\boldsymbol{I}$ 是单位矩阵。然后最小二乘法估计 $\hat{\boldsymbol{\theta}}$ 具有以下性质：

1）$\hat{\boldsymbol{\theta}}$ 为 θ_0 的无偏估计。

2）$\hat{\boldsymbol{\theta}}$ 的协方差矩阵由下式给出：

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}})=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}$$

证明：对于估计 $\hat{\boldsymbol{\theta}}$ 是无偏的，它必须满足

$$E[\hat{\boldsymbol{\theta}}]=\theta_0 \tag{4.8}$$

式中，θ_0 是真参数。

通过采用期望

$$\begin{aligned}E[\hat{\boldsymbol{\theta}}]&=E[(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}]\\&=E[(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\theta_0+\boldsymbol{v})]\\&=E[(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})\theta_0]+E[(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{v}]\\&=E[\theta_0]+(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}E[\boldsymbol{v}]\\&=\theta_0\end{aligned} \tag{4.9}$$

计算出 $\hat{\boldsymbol{\theta}}$ 为 θ_0 的无偏估计。

$\hat{\boldsymbol{\theta}}$ 的协方差矩阵由下式确定：

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}})=E[(\hat{\boldsymbol{\theta}}-E[\hat{\boldsymbol{\theta}}])(\hat{\boldsymbol{\theta}}-E[\hat{\boldsymbol{\theta}}])^{\mathrm{T}}] \tag{4.10}$$

式中，$\hat{\boldsymbol{\theta}}=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}$ 由式（4.9）和式（4.10）可写作

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol{\theta}})&=E[((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}-\theta_0)((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}-\theta_0)^{\mathrm{T}}]\\&=E[((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\theta_0+v)-\theta_0)((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\theta_0+v)-\theta_0)^{\mathrm{T}}]\\&=E[((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}v)((\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}v)^{\mathrm{T}}]\\&=E[(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}vv^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}]\\&=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}E[vv^{\mathrm{T}}]\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\\&=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\end{aligned} \tag{4.11}$$

证毕。

给出估算：

$$\hat{\boldsymbol{\theta}}_w=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{Y}$$

被称为 $\boldsymbol{\theta}$ 的最佳线性无偏估计。它也被称为加权最小二乘估计，并具有下列性质：

1）$\hat{\boldsymbol{\theta}}_w$ 是 $\boldsymbol{\theta}$ 的无偏估计。

2）$\hat{\boldsymbol{\theta}}_w$ 的协方差矩阵由下式给出：

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}_w)=(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\Phi})^{-1}$$

3）$\hat{\boldsymbol{\theta}}_w$ 是定义为 $\hat{\boldsymbol{\theta}}_{\mathrm{linear}}=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Y}$ 的线性估计中的最小方差估计，其中 $\boldsymbol{Z}$ 是一个 $\boldsymbol{N}\times n$ 的常数矩阵。

详细的证明由 Söderström 和 Stoica（1989）给出。

在这里，将在下面的方程式中给出通过等效“白化”步骤导出最佳线性无偏估计。

回归模型由下式给出：

$$\boldsymbol{Y}=\boldsymbol{\Phi}\boldsymbol{\theta}+v \tag{4.12}$$

式中，$\mathrm{Cov}(v)=\boldsymbol{R}$ 不是对角矩阵。

等式两边乘以 $\boldsymbol{R}^{-1/2}$ 得出

$$\underbrace{\boldsymbol{R}^{-1/2}\boldsymbol{Y}}_{\boldsymbol{Y}_f}=\underbrace{\boldsymbol{R}^{-1/2}\boldsymbol{\Phi}\boldsymbol{\theta}}_{\boldsymbol{\Phi}_f}+\underbrace{\boldsymbol{R}^{-1/2}v}_{v_f}$$

v_f 的协方差矩阵可表示如下：

$$\begin{aligned}\mathrm{Cov}(v_f)&=\mathrm{Cov}(\boldsymbol{R}^{-1/2}v)\\&=E[(\boldsymbol{R}^{-1/2}v)(\boldsymbol{R}^{-1/2}v)^{\mathrm{T}}]\\&=E[\boldsymbol{R}^{-1/2}vv^{\mathrm{T}}\boldsymbol{R}^{-1/2}]\\&=\boldsymbol{R}^{-1/2}E[vv^{\mathrm{T}}]\boldsymbol{R}^{-1/2}\\&=\boldsymbol{R}^{-1/2}\boldsymbol{R}\boldsymbol{R}^{-1/2}=\boldsymbol{I}\end{aligned} \tag{4.13}$$

式（4.13）表明 v_f 由白噪声组成，所以上述过程等效于白化噪声 v_t。由于 v_f 现

在由白噪声组成，可以应用普通最小二乘法来估计模型参数 $\boldsymbol{\theta}$，即

$$\begin{aligned}\hat{\boldsymbol{\theta}}_w &= (\boldsymbol{\Phi}_f^{\mathrm{T}}\boldsymbol{\Phi}_f)^{-1}\boldsymbol{\Phi}_f^{\mathrm{T}}\boldsymbol{Y}_f \\ &= (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{Y}\end{aligned}$$

在普通最小二乘估计得出的基础结果上，协方差矩阵 $\hat{\boldsymbol{\theta}}_w$ 由下式给出：

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol{\theta}}_w) &= (\boldsymbol{\Phi}_f^{\mathrm{T}}\boldsymbol{\Phi}_f)^{-1} \\ &= ((\boldsymbol{R}^{-1/2}\boldsymbol{\Phi})^{\mathrm{T}}\boldsymbol{R}^{-1/2}\boldsymbol{\Phi})^{-1} \\ &= (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1/2}\boldsymbol{R}^{-1/2}\boldsymbol{\Phi})^{-1} \\ &= (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\Phi})^{-1}\end{aligned} \tag{4.14}$$

例 4.1.1　离散时间传递函数模型由下式给出：

$$y_t = \frac{b_{10} + b_{20}z^{-1}}{1 + a_{10}z^{-1}}u_{t-d} + (1 + c_{10}z^{-1})e_t \tag{4.15}$$

如果我们用加权最小二乘法来解决估计问题，求得到的加权矩阵。

离散时间传递函数模型可以被变换为

$$(1 + a_{10}z^{-1})y_t = (b_{10} + b_{20}z^{-1})u_{t-d} + (1 + a_{10}z^{-1})(1 + c_{10}z^{-1})e_t \tag{4.16}$$

变形为

$$y_t = -a_{10}y_{t-1} + b_{10}u_{t-d} + b_{20}u_{t-d-1} + v_t \tag{4.17}$$

式中

$$v_t = e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2} \tag{4.18}$$

令

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \tag{4.19}$$

构造对应的协方差矩阵：

$$\boldsymbol{R} = \begin{pmatrix} r_v(0) & r_v(1) & \cdots & r_v(n-1) \\ r_v(-1) & r_v(0) & \cdots & r_v(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ r_v(1-n) & r_v(2-n) & \cdots & r_v(0) \end{pmatrix} \tag{4.20}$$

接下来，我们需要评估当 τ 取不同值时的协方差函数 $r_v(\tau)$。

$$\begin{aligned}r_v(0) &= E[v_t v_t] \\ &= E[(e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2})(e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2})] \\ &= \lambda^2(1 + (a_{10} + c_{10})^2 + a_{10}^2c_{10}^2)\end{aligned} \tag{4.21}$$

$$\begin{aligned}r_v(1) &= E[v_{t+1}v_t] \\ &= E[(e_{t+1} + (a_{10} + c_{10})e_t + a_{10}c_{10}e_{t-1})(e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2})] \\ &= \lambda^2(a_{10} + c_{10} + a_{10}c_{10}(a_{10} + c_{10}))\end{aligned} \tag{4.22}$$

$$r_v(2) = E[v_{t+2}v_t]$$

$$= E[(e_{t+2} + (a_{10} + c_{10})e_{t+1} + a_{10}c_{10}e_t)(e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2})]$$
$$= \lambda^2 a_{10}c_{10} \tag{4.23}$$

同时，$r_v(\tau) = 0$，$\tau \geqslant 3$。令 $\alpha = a_{10} + c_{10}$ 和 $\beta = a_{10}c_{10}$ 使

$$\boldsymbol{R} = \lambda^2 \begin{pmatrix} 1+\alpha^2+\beta^2 & \alpha(1+\beta) & \beta & & & \\ \alpha(1+\beta) & 1+\alpha^2+\beta^2 & \alpha(1+\beta) & \ddots & & \\ \beta & \alpha(1+\beta) & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \beta \\ & & \ddots & \ddots & \ddots & \alpha(1+\beta) \\ & & & \beta & \alpha(1+\beta) & 1+\alpha^2+\beta^2 \end{pmatrix} \tag{4.24}$$

4.2 预测误差方法

显然，并不是所有的参数估计问题都可以通过线性回归方法来解决。下面的例子说明，需要比线性回归先进的辨识方法。

例 4.2.1 再考虑例 4.1.1。讨论普通最小二乘回归是否以这种模型能否提供一致估计。

原方程变形得到另一种形式

$$(1 + a_{10}z^{-1})y_t = (b_{10} + b_{20}z^{-1})u_{t-d} + (1 + a_{10}z^{-1})(1 + c_{10}z^{-1})e_t \tag{4.25}$$

模型可变形得到

$$y_t = -a_{10}y_{t-1} + b_{10}u_{t-d} + b_{20}u_{t-d-1} + v_t \tag{4.26}$$

式中

$$v_t = e_t + (a_{10} + c_{10})e_{t-1} + a_{10}c_{10}e_{t-2} \tag{4.27}$$

定义下列矢量

$$\boldsymbol{\varphi}_t = \begin{pmatrix} -y_{t-1} \\ u_{t-d} \\ u_{t-d-1} \end{pmatrix} \qquad \boldsymbol{\theta}_0 = \begin{pmatrix} a_{10} \\ b_{10} \\ b_{20} \end{pmatrix} \tag{4.28}$$

然后，回归模型对于式（4.26）变为

$$\boldsymbol{y}_t = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v_t \tag{4.29}$$

而实际过程模型是

$$\boldsymbol{y}_t = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}_0 + \boldsymbol{v}_t \tag{4.30}$$

对于普通最小二乘回归 $\boldsymbol{\theta} = (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{Y}$，有

$$\boldsymbol{\Phi} = \begin{pmatrix} \varphi^{\mathrm{T}}(1) \\ \varphi^{\mathrm{T}}(2) \\ \vdots \\ \varphi^{\mathrm{T}}(N) \end{pmatrix} \tag{4.31}$$

从而

$$\hat{\boldsymbol{\theta}} = \left(\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right)^{-1}\left(\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{y}_t\right)$$
$$= \left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{y}_t\right) \tag{4.32}$$

由式（4.30）替换 y_t 得

$$\hat{\boldsymbol{\theta}} = \left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)(\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}_0 + \boldsymbol{v}_t)\right)$$
$$= \boldsymbol{\theta}_0 + \left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\,\boldsymbol{v}_t\right) \tag{4.33}$$

证明 $\left(\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right)^{-1}$存在，也就是说，在足够的阶数输入持续地激励。当 $N\to\infty$ 时有

$$\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) \to E[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)] \tag{4.34}$$

$$\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\,\boldsymbol{v}_t \to E[\boldsymbol{\varphi}(t)\,\boldsymbol{v}_t] \tag{4.35}$$

从而

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + (E[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)])^{-1}E[\boldsymbol{\varphi}(t)\boldsymbol{v}_t] \tag{4.36}$$

$\hat{\boldsymbol{\theta}}$ 的一致性取决于是否 $E[\boldsymbol{\varphi}(t)\,\boldsymbol{v}_t]=0$。可以证明：

$$E[\boldsymbol{\varphi}(t)\boldsymbol{v}_t] = E\begin{pmatrix} -y_{t-1}v_t \\ u_{t-d}v_t \\ u_{t-d-1}v_t \end{pmatrix} \neq 0 \tag{4.37}$$

它是直接表明 $E[u_{t-d}\,v_t]=0$ 且 $E[u_{t-d-1}\,v_t]=0$。然而，可以验证 $E[-y_t-1v_t]\neq 0$。最后，可以得出结论，如果应用普通最小二乘法，此估计不具有一致性。

4.2.1　最优预测

当普通最小二乘回归不能提供一致的估计时，为了找到合适的辨识算法，需要回顾辨识的目的和过程。

辨识是要找到一组基于在给定的输入和输出中预测与实际测量值尽可能接近的参数。该预测是由根据预示变量来的之前数据确定的。该预测是从模型中得出的。辨识过程的第一步是得出一个最佳预示变量，可以最大限度地减少预测误差。

预测误差由下式确定：

$$\varepsilon(t,\theta) = y_t - \hat{y}(t\,|\,t-1)$$

式中，$\hat{y}(t\,|\,t-1)$ 表示对 y_t 的预测，其中给定所有的数据采样及时刻 $t-1$（如 y_{t-1}，u_{t-1}，y_{t-2}，u_{t-2}，…）。

为了说明这一点，考虑 ARX 模型：

$$y_t = -a_1 y_{t-1} - a_2 y_{t-2} - \cdots - a_{na} y_{t-na} + b_1 u_{t-1} + b_2 u_{t-2} + \cdots + b_{nb} u_{t-nb} + e_t$$

选择以下预示变量作为提前预测的 y_t，基于所有在时间上的已知信息 $t-1$，即 y_{t-1}，…，y_{t-na}，和 u_{t-1}，…，u_{t-nb}：

$$\hat{y}(t|t-1) = -a_1 y_{t-1} - \cdots - a_{na} y_{t-na} + b_1 u_{t-1} + \cdots + b_{nb} u_{t-nb}$$

显然，这个预示变量也是回归方程，被用于估计 ARX 模型，也就是

$$\hat{y}(t|t-1) = \varphi^{\mathrm{T}}(t)\theta$$

如果已知具体模型参数，那么预测误差为

$$y_t - \hat{y}(t|t-1) = e_t$$

即白噪声。直观地看，如果一个预测，有白噪声作为其预测误差应该是最优的预测因子，然而白噪声不能被进一步预测，将在这个案例中马上看出。

回想一下，损失函数为 ARX 模型由下式给出：

$$\boldsymbol{J}(\theta) = \frac{1}{N}(\boldsymbol{Y} - \boldsymbol{\Phi\theta})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{\Phi\theta})$$

可进一步写为

$$J(\theta) = \frac{1}{N}\sum_{t=1}^{N}(y_t - \varphi^{\mathrm{T}}(t)\theta)^2 = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}(t|t-1))^2$$

这表明普通最小二乘回归的 ARX 模型的目标是使预测误差最小化，只有当预测器是最佳时才能实现。现在，一个有趣的问题是，是否可以得到最优的预测进程，这不能由一个 ARX 模型进行说明。考虑到一般线性模型：

$$y_t = G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t \tag{4.38}$$

假设 $G_p(0;\ \theta)=0$。请注意，所有之前的线性模型，如 ARX、ARMAX、BJ（Box-Jenkins）、OE（输出误差）和 PEM（预测误差法）这是一般线性模型的特殊情况。一般线性的之前一步预测已经由 Söderström 和 Stoica（1989）描述，即

$$\hat{y}(t|t-1) = L_1(z^{-1};\theta)y_t + L_2(z^{-1};\theta)u_t \tag{4.39}$$

作为预测必须依靠过去的输出数据，该滤波器由下式限制

$$L_1(0;\theta) = 0 \tag{4.40}$$

从式（4.38）减去式（4.39）得出以下预测误差：

$$\begin{aligned}\varepsilon(t,\theta) &= G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t - L_1(z^{-1};\theta)y_t - L_2(z^{-1};\theta)u_t \\ &= G_p(z^{-1};\theta)u_t + (G_l(z^{-1};\theta) - I)e_t + e_t - L_1(z^{-1};\theta)y_t - L_2(z^{-1};\theta)u_t \\ &= (G_p(z^{-1};\theta) - L_2(z^{-1};\theta))u_t + (G_l(z^{-1};\theta) - I)e_t - L_1(z^{-1};\theta)y_t + e_t\end{aligned}$$

从式（4.38），e_t 可以推导为

$$e_t = G_l^{-1}(z^{-1};\theta)(y_t - G_p(z^{-1};\theta)u_t)$$

利用该关系，可以进一步写出预测误差的表达式为

$$\begin{aligned}\varepsilon(t,\theta) = {} & (G_p(z^{-1};\ \theta) - L_2(z^{-1};\theta))u_t + (G_l(z^{-1};\theta) - I) \\ & \times G_l^{-1}(z^{-1};\theta)(y_t - G_p(z^{-1};\theta)u_t) - L_1(z^{-1};\theta)y_t + e_t\end{aligned}$$

$$= \underbrace{(G_l^{-1}(z^{-1};\theta)G_p(z^{-1};\theta) - L_2(z^{-1};\theta))}_{\Psi_u(z^{-1};\theta)}u_t + \underbrace{(I - G_l^{-1}(z^{-1};\theta) - L_1(z^{-1};\theta))}_{\Psi_y(z^{-1};\theta)}y_t + e_t$$

$$= \Psi_u(z^{-1};\theta)u_t + \Psi_y(z^{-1};\theta)y_t + e_t$$

给出条件 $G_l(0;\ \theta) = I$ 和 $L_1(0;\ \theta) = 0$，可以验证

$$\Psi_y(0;\theta) = 0$$

即 $\Psi_y\ (z^{-1};\ \theta)$ 具有至少一个样本时间延迟。因此，通过扩大传递函数变换为脉冲响应的形式，将有

$$\Psi_y(z^{-1};\theta)y_t = \psi_{y1}y_{t-1} + \psi_{y2}y_{t-2} + \cdots$$

注意到 u_t 是独立于 e_t 的，并作为关于 $\Psi_y(z^{-1};\ \theta)y_t$ 的未来白噪声干扰，e_t 是独立于 $\Psi_u(z^{-1};\ \theta)$ 和 $\Psi_y(z^{-1};\ \theta)y_t$ 的，所以

$$\mathrm{Cov}(\boldsymbol{\varepsilon}(t,\theta)) = \mathrm{Cov}[\Psi_u(z^{-1};\theta)u_t + \Psi_y(z^{-1};\theta)y_t] + \mathrm{Cov}[e_t] \geqslant \mathrm{Cov}(e_t)$$

或者

$$\mathrm{trace}[\mathrm{Cov}(\boldsymbol{\varepsilon}(t,\theta))] \geqslant \mathrm{trace}[\mathrm{Cov}(e_t)]$$

显然，$[\mathrm{Cov}(\boldsymbol{\varepsilon}(t,\ \theta))]$ 轨迹的下限即 $[\mathrm{Cov}(e_t)]$ 的轨迹，也就是说，白噪声方差的轨迹是 e_t，它被定义为 Σ_e 因此，最佳的前一步预测应该给这个下限作为其预测误差。这通过下式实现：

$$\Psi_u\ (z^{-1};\ \theta) = 0$$

$$\Psi_y\ (z^{-1};\ \theta) = 0$$

分别解出这两个方程：

$$L_2(z^{-1};\theta) = G_l^{-1}(z^{-1};\theta)G_p(z^{-1};\theta)$$

$$L_1(z^{-1};\theta) = I - G_l^{-1}(z^{-1};\theta)$$

结果得到最佳预测。最佳预测误差是 $\boldsymbol{\varepsilon}(t,\ \theta) = e_t$，即白噪声。这个结果并不令人惊讶，因为白噪声不包含任何可预见的内容。如果预测能够给白噪声作为其预测误差，已经应该实现了最好的预测。因此，确定一个单步向前预测器是否已达到其最佳，可以检查其预测误差是否是白噪声。概括地说，对于一般的线性模型

$$y_t = G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t \tag{4.41}$$

最佳的预测是

$$\begin{aligned}\hat{y}(t\,|\,t-1) &= L_1(z^{-1};\theta)y_t + L_2(z^{-1};\theta)u_t \\ &= G_l^{-1}(z^{-1};\theta)G_p(z^{-1};\theta)u_t + [I - G_l^{-1}(z^{-1};\theta)]y_t \end{aligned} \tag{4.42}$$

与最佳单步向前预测，将预测误差表示为

$$\boldsymbol{\varepsilon}(t,\theta) = G_l^{-1}(z^{-1};\theta)[y_t - G_p(z^{-1};\theta)u_t] \tag{4.43}$$

当模型 $G_p(z^{-1};\ \theta)$ 和 $G_l(z^{-1};\ \theta)$ 为最佳时，预测误差就是白噪声 e_t。

对于状态空间模型：

$$x_{t+1} = \boldsymbol{A}(\theta)x_t + \boldsymbol{B}(\theta)u_t + w_t$$
$$y_t = \boldsymbol{C}(\theta)x_t + e_t$$

式中，w_t 和 e_t 是具有零期望和协方差矩阵 $\boldsymbol{R}_1(\theta)$ 和 $\boldsymbol{R}_2(\theta)$ 相互不相关的白噪声序列，最优的预测值是卡尔曼预测器：

$$\hat{x}(t+1|t) = \boldsymbol{A}(\theta)\hat{x}(t|t-1) + \boldsymbol{B}(\theta)u_t + \boldsymbol{K}(\theta)[y_t - \boldsymbol{C}(\theta)\hat{x}(t|t-1)]$$
$$\hat{y}(t|t-1) = \boldsymbol{C}(\theta)\hat{x}(t|t-1)$$

卡尔曼增益由下式给出：

$$\boldsymbol{K}(\theta) = \boldsymbol{A}(\theta)\boldsymbol{P}(\theta)\boldsymbol{C}^{\mathrm{T}}(\theta)[\boldsymbol{C}(\theta)\boldsymbol{P}(\theta)\boldsymbol{C}^{\mathrm{T}}(\theta) + \boldsymbol{R}_2(\theta)]^{-1}$$

式中，$\boldsymbol{P}(\theta)$ 是下列 Riccati 代数方程的解：

$$\boldsymbol{P}(\theta) = \boldsymbol{A}(\theta)\boldsymbol{P}(\theta)\boldsymbol{A}^{\mathrm{T}}(\theta) + \boldsymbol{R}_1(\theta) - \boldsymbol{K}(\theta)\boldsymbol{C}(\theta)\boldsymbol{P}(\theta)\boldsymbol{A}^{\mathrm{T}}(\theta)$$

其结果是，卡尔曼滤波器形式（也称为创新形式）的状态空间模型通常用于系统识别的状态空间模型结构。

例 4.2.2　再考虑例 4.1.1。

1）得出最优的一步超前预测。讨论为什么它是最优的。

2）得出最优的两步超前预测，并讨论为什么它给出了一个最优的两步超前预测。

白噪声项 e_t 可由式（4.25）推导出：

$$e_t = \frac{1}{1+c_{10}z^{-1}}y_t - \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{(1+a_{10}z^{-1})(1+c_{10}z^{-1})}u_t \tag{4.44}$$

同时，式（4.25）可以改写为

$$y_t = \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}u_t + c_{10}z^{-1}e_t + e_t \tag{4.45}$$

用式（4.44）展开式代替式（4.45）的倒数第二项 e_t：

$$\begin{aligned}
y_t &= \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}u_t + c_{10}z^{-1}\left[\frac{1}{1+c_{10}z^{-1}}y_t - \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{(1+a_{10}z^{-1})(1+c_{10}z^{-1})}u_t\right] + e_t \\
&= \frac{c_{10}z^{-1}}{1+c_{10}z^{-1}}y_t + \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}\left[1 - \frac{c_{10}z^{-1}}{1+c_{10}z^{-1}}\right]u_t + e_t \\
&= \frac{c_{10}z^{-1}}{1+c_{10}z^{-1}}y_t + \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{(1+a_{10}z^{-1})(1+c_{10}z^{-1})}u_t + e_t
\end{aligned} \tag{4.46}$$

使一步超前预测为

$$\begin{aligned}
\hat{y}(t|t-1) &= \frac{c_{10}z^{-1}}{1+c_{10}z^{-1}}y_t + \frac{(b_{10}+b_{20}z^{-1})z^{-d}}{(1+a_{10}z^{-1})(1+c_{10}z^{-1})}u_t \\
&= L_1(z^{-1};\theta)y_t + L_2(z^{-1};\theta)u_t
\end{aligned} \tag{4.47}$$

满足 $L_1(0;\theta) = 0$。

对此预测，将预测误差由下式给出：

$$\varepsilon(t,\theta) = y_t - \hat{y}(t|t-1) = e_t \tag{4.48}$$

由于这是白噪声，因而，此一步超前预测以式（4.47）表示是最优的。

一般线性两步超前预测可描述为

$$\hat{y}(t|t-2)=L_1(z^{-1};\theta)y_t+L_2(z^{-1};\theta)u_t \tag{4.49}$$

式中，L_1（z^{-1}；θ）具有至少两个样本延迟才成为一个两步超前预测。

预测误差表示为

$$\begin{aligned}\varepsilon(t,\theta)&=y_t-\hat{y}(t|t-2)\\&=\frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}u_t+c_{10}z^{-1}e_t+e_t-L_1(z^{-1};\theta)y_t-L_2(z^{-1};\theta)u_t\\&=-L_1(z^{-1};\theta)y_t+\left(\frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}-L_2(z^{-1};\theta)\right)u_t+c_{10}z^{-1}e_t+e_t\end{aligned} \tag{4.50}$$

$L_1(z^{-1}$；θ）具有至少两个采样的延迟，预测误差的方差为

$$\begin{aligned}\mathrm{Var}(\varepsilon(t,\theta))&=\mathrm{Var}\left(-L_1(z^{-1};\theta)y_t+\left(\frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}-L_2(z^{-1};\theta)\right)u_t\right)\\&\quad+\mathrm{Var}(c_{10}z^{-1}e_t+e_t)\\&\geqslant\mathrm{Var}(c_{10}z^{-1}e_t+e_t)=\lambda^2(c_{10}{}^2+1)\end{aligned} \tag{4.51}$$

鉴于式（4.51）和式（4.49），最小方差 λ^2（c_{10}^2+1）实现了与以下两步超前最优预测：

$$\hat{y}(t|t-2)=\frac{(b_{10}+b_{20}z^{-1})z^{-d}}{1+a_{10}z^{-1}}u_t \tag{4.52}$$

因此，预测误差给定为

$$\varepsilon(t,\theta)=y_t-\hat{y}(t|t-2)=c_{10}z^{-1}e_t+e_t=e_t+c_{10}e_{t-1} \tag{4.53}$$

4.2.2　预测误差方法

预测误差的方法是为了找到 $\hat{\theta}$，使得预测误差 $\varepsilon(1,\ \theta)$，$\varepsilon(2,\ \theta)$，…，$\varepsilon(N,\ \theta)$最小化。不像设计不同的最佳预测指标，其中精确的模型参数被假定为已知的，在系统辨识的模型中参数是未知的。因此，预测将是未知参数的函数。为限定预测对于系统辨识是有用的，预测值和模型之间的关系应该是唯一的。当预测值被修正后，模型参数应该唯一确定。所选择的预测器是最优的预测，因为只有最佳预测器可以给预测误差最小。作为结果，识别问题可以转化为一个优化问题，并且参数估计可以继承某些最优性。

考虑以下预测误差的目标函数：

$$J_N(\theta)=\mathrm{trace}[\Sigma_N(\theta)]\triangleq\mathrm{trace}\left[\frac{1}{N}\sum_{t=1}^{N}\varepsilon(t,\theta)\varepsilon^{\mathrm{T}}(t,\theta)\right]$$

考虑到实际系统由下式给出：

$$y_t=G_p(z^{-1};\theta_0)u_t+G_l(z^{-1};\theta_0)e_t \tag{4.54}$$

式（4.43）代入式（4.54）得出

$$\begin{aligned}\varepsilon(t,\theta) &= G_l^{-1}(z^{-1};\theta)[y_t - G_p(z^{-1};\theta)u_t] \\ &= G_l^{-1}(z^{-1};\theta)[G_p(z^{-1};\theta_0) - G_p(z^{-1};\theta_0)]u_t + G_l^{-1}(z^{-1};\theta)G_l(z^{-1};\theta_0)e_t \\ &\triangleq \Phi_u(z^{-1};\theta,\theta_0)u_t + \Phi_e(z^{-1};\theta,\theta_0)e_t \\ &= \Phi_u(z^{-1};\theta,\theta_0)u_t + (\Phi_e(z^{-1};\theta,\theta_0) - I)e_t + e_t\end{aligned} \tag{4.55}$$

式中

$$\Phi_u(z^{-1};\theta,\theta_0) = G_l^{-1}(z^{-1};\theta)[G_p(z^{-1};\theta_0) - G_p(z^{-1};\theta)] \tag{4.56}$$

$$\Phi_e(z^{-1};\theta,\theta_0) = G_l^{-1}(z^{-1};\theta)G_l(z^{-1};\theta_0) \tag{4.57}$$

通过对整个过程和干扰模型的假设，$G_p(0;\ \theta)=0$、$G_l(0;\ \theta)=I$、$G_p(0;\ \theta_0)=0$ 和 G_l（0；θ_0）$=I$，它可以从式（4.56）和式（4.57）进行验证：

$$\Phi_u(0;\theta,\theta_0) = 0 \tag{4.58}$$

$$\Phi_e(0;\theta,\theta_0) - I = 0 \tag{4.59}$$

式（4.58）和式（4.59）意味着Φ_u（0，θ，θ_0）及（Φ_e(0；θ，θ_0）$-I$）应至少有一个样本的时间延迟。需要注意的是，u_t 和 e_t 是独立的，因此

$$\mathrm{Cov}[\varepsilon(t,\theta)] = \mathrm{Cov}[\Phi_u(z^{-1};\theta,\theta_0)u_t] + \mathrm{Cov}[(\Phi_e(z^{-1};\theta,\theta_0) - I)e_t] + \mathrm{Cov}[e_t] \geqslant \mathrm{Cov}[e_t]$$

这得到

$$\mathrm{trace}[\mathrm{Cov}(\varepsilon(t,\theta))] \geqslant \mathrm{trace}[\mathrm{Cov}(e_t)]$$

再一次，缩小范围是白噪声 Σ_e 协方差的轨迹。当为如下时其为最小值：

$$\Phi_u(z^{-1};\theta,\theta_0) = 0 \tag{4.60}$$

$$\Phi_e(z^{-1};\theta,\theta_0) = I \tag{4.61}$$

比较式（4.60）和式（4.61）与式（4.56）和式（4.57），达到下界意味着估计的一致性，即

$$G_p(z^{-1};\theta) = G_p(z^{-1};\theta_0) \tag{4.62}$$

$$G_l(z^{-1};\theta) = G_l(z^{-1};\theta_0) \tag{4.63}$$

PEM 的算法可概述如下：

1）选择一个模型结构：

$$y_t = G_p(z^{-1};\theta)u_t + G_l(z^{-1};\theta)e_t$$

并写出相应的最优预测：

$$\hat{y}(t|t-1) = G_l^{-1}(z^{-1};\theta)G_p(z^{-1};\theta)u_t + [I - G_l^{-1}(z^{-1};\theta)]y_t$$

2）替代实验数据，u_1，u_2，…，u_N 和 y_1，y_2，…，y_N，最佳预测得到单步向前预测 $\hat{y}(1|0)$，$\hat{y}(2|1)$，…，$\hat{y}(N|N-1)$，均为 θ 的函数。对应的预测误差然后可以由测得的输出和预测 $\varepsilon(1,\ \theta)$，$\varepsilon(2,\ \theta)$，…，$\varepsilon(N,\ \theta)$ 确定。

3）确定的预测误差的采样协方差为：

$$\Sigma_N(\theta) = \frac{1}{N-p}\sum_{t=1}^{N}\varepsilon(t,\theta)\varepsilon^{\mathrm{T}}(t,\theta)$$

式中，p 为估计参数的数值。

通过迹减少 $\Sigma_N(\theta)$ 来求 $\hat{\theta}$。最小二乘问题一般是非线性的和数值解，这通常是必要的。

当 $N\to\infty$ 时，$\Sigma_N(\theta)\to\mathrm{Cov}[\varepsilon(t,\theta)]$。如果上述优化过程实现了全局最优解，将得到白色噪声的预测误差，然后将估计的参数会收敛到真实参数。因此，该预测对一般线性模型误差方法能够实现一致估计，即 $\hat{\theta}\xrightarrow{N\to\infty}\theta_0$。

例 4.2.3　考虑再次给出前面例子中的真正离散时间模型：

$$y_t=\frac{b_{10}+b_{20}z^{-1}}{1+a_{10}z^{-1}}u_{t-d}+(1+c_{10}z^{-1})e_t$$

1）如果模型被选择为具有相同的结构作为该真正模型，表明该 PEM 算法给出了这两种过程和扰动模型的一致估计。

2）如果模型被选择为具有相同的结构作为该真正模型，讨论最佳两步向前预测误差的最小化是否会提供该过程的一致性估计和噪声模型。

最优单步向前预测和相应的预测误差可以推导如下：

$$\begin{aligned}\hat{y}(t|t-1)&=G_l^{-1}(z^{-1};\theta)G_p(z^{-1};\theta)u_t+[I-G_l^{-1}(z^{-1};\theta)]y_t\\&=\frac{c_1z^{-1}}{1+c_1z^{-1}}y_t+\frac{z^{-d}(b_1+b_2z^{-1})}{(1+a_1z^{-1})(1+c_1z^{-1})}u_t\end{aligned}\tag{4.64}$$

$$\begin{aligned}\varepsilon(t,\theta)&=G_l^{-1}(z^{-1};\theta)[y_t-G_p(z^{-1};\theta)u_t]\\&=\frac{1}{1+c_1z^{-1}}\left[y_t-\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}}u_t\right]\\&=\frac{1}{1+c_1z^{-1}}\left[\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{1+a_{10}z^{-1}}u_t+(1+c_{10}z^{-1})e_t-\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}}u_t\right]\\&=\left[\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{(1+c_1z^{-1})(1+a_{10}z^{-1})}-\frac{z^{-d}(b_1+b_2z^{-1})}{(1+c_1z^{-1})(1+a_1z^{-1})}\right]u_t+\frac{1+c_{10}z^{-1}}{1+c_1z^{-1}}e_t\\&=\Phi_u(z^{-1};\theta,\theta_0)u_t+\Phi_e(z^{-1};\theta,\theta_0)e_t\\&=\Phi_u(z^{-1};\theta,\theta_0)u_t+(\Phi_e(z^{-1};\theta,\theta_0)-I)e_t+e_t\end{aligned}\tag{4.65}$$

$\Phi_u(z^{-1};\theta,\theta_0)$ 和 $\Phi_e(z^{-1};\theta,\theta_0)$ 至少一个具有采样时间延迟，它可以验证：

$$\Phi_u(0;\theta,\theta_0)=0$$

$$\Phi_e(0;\theta,\theta_0)-I=0$$

如果 u_t 和 e_t 是独立的，它可以证明：

$$\begin{aligned}\mathrm{Var}[\varepsilon(t,\theta)]&=\mathrm{Var}[\Phi_u(z^{-1};\theta,\theta_0)u_t]+\mathrm{Var}[(\Phi_e(z^{-1};\theta,\theta_0)-I)e_t]\\&\quad+\mathrm{Var}[e_t]\geqslant\mathrm{Var}[e_t]\end{aligned}\tag{4.66}$$

达到最小时，当

$$\Phi_u(z^{-1};\theta,\theta_0)=0\tag{4.67}$$

$$\Phi_e(z^{-1};\theta,\theta_0)=I\tag{4.68}$$

实现最小值意味着

$$\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{(1+c_1z^{-1})(1+a_{10}z^{-1})}=\frac{z^{-d}(b_1+b_2z^{-1})}{(1+c_1z^{-1})(1+a_1z^{-1})} \tag{4.69}$$

$$1+c_{10}z^{-1}=1+c_1z^{-1} \tag{4.70}$$

也就是说，一致性的实现。

类似于式（4.52）的推导，两个向前一步最优预测可以推导出：

$$\hat{y}(t|t-2)=\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}}u_t \tag{4.71}$$

因此，预测误差由下式得出：

$$\begin{aligned}\varepsilon(t,\theta)&=y_t-\hat{y}(t|t-2)\\&=\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{1+a_{10}z^{-1}}u_t+(1+c_{10}z^{-1})e_t-\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}}u_t\\&=\left[\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{1+a_{10}z^{-1}}-\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}}\right]u_t+(1+c_{10}z^{-1})e_t\\&=\Phi_u(z^{-1};\theta,\theta_0)u_t+(1+c_{10}z^{-1})e_t\end{aligned} \tag{4.72}$$

当 u_t 和 e_t 独立时，表示如下：

$$\begin{aligned}\mathrm{Var}[\varepsilon(t,\theta)]&=\mathrm{Var}[\Phi_u(z^{-1};\theta,\theta_0)u_t]+\mathrm{Var}[(1+c_{10}z^{-1})e_t]\\&\geqslant\mathrm{Var}[(1+c_{10}z^{-1})e_t]\end{aligned} \tag{4.73}$$

达到最小时当

$$\Phi_u(z^{-1};\theta,\theta_0)=0 \tag{4.74}$$

实现最小值意味着

$$\frac{z^{-d}(b_{10}+b_{20}z^{-1})}{1+a_{10}z^{-1}}=\frac{z^{-d}(b_1+b_2z^{-1})}{1+a_1z^{-1}} \tag{4.75}$$

也就是说，这是一致性实现的过程模型，但对噪声模型没有什么影响。

4.2.3 独立参数预测误差法

考虑过程和干扰模型被独立参数化的一般线性模型：

$$y_t=G_p(z^{-1},\theta_1)u_t+G_l(z^{-1},\theta_2)e_t \tag{4.76}$$

而真正的模型由下式给出：

$$y_t=G_p(z^{-1},\theta_{10})u_t+G_l(z^{-1},\theta_{20})e_t \tag{4.77}$$

最优单步向前预测是与上述同源，除了对于不同的参数设置：

$$\hat{y}(t|t-1)=G_l^{-1}(z^{-1},\theta_2)G_p(z^{-1},\theta_1)u_t+[I-G_l^{-1}(z^{-1},\theta_2)]y_t \tag{4.78}$$

预测误差由下式给出：

$$\begin{aligned}\varepsilon(t,\theta_1,\theta_2)&=G_l^{-1}(z^{-1};\theta_2)[y_t-G_p(z^{-1};\theta_2)u_t]\\&=G_l^{-1}(z^{-1};\theta_2)[G_p(z^{-1};\theta_{10})-G_p(z^{-1};\theta_1)]u_t\\&\quad+G_l^{-1}(z^{-1};\theta_2)G_l(z^{-1};\theta_{20})e_t\end{aligned}$$

$$\triangleq \Phi_u(z^{-1},\theta_1,\theta_{10},\theta_2)u_t + \Phi_e(z^{-1};\theta_2,\theta_{20})e_t \tag{4.79}$$

由于 u_t 和 e_t 是独立的，对式（4.79）两边同时进行操作得出

$$\mathrm{Cov}[\varepsilon(t,\theta_1,\theta_2)] = \mathrm{Cov}[\Phi_u(z^{-1};\theta_1,\theta_{10},\theta_2)u_t] + \mathrm{Cov}[\Phi_e(z^{-1};\theta_2,\theta_{20})e_t] \tag{4.80}$$

式（4.80）中尽量减少关于 θ_1 的量得到以下关系式：

$$\mathrm{argmin}\theta_1\mathrm{Cov}[\varepsilon(t,\theta_1,\theta_2)] = \mathrm{argmin}\theta_1\mathrm{Cov}[\Phi_u(z^{-1};\theta_1,\theta_{10},\theta_2)u_t]$$

当 $\Phi_u(z^{-1};\ \theta_1,\ \theta_{10},\ \theta_2)=0$ 时，得到最小值。这意味着 G_p（z^{-1}，θ_1）$=G_p$（z^{-1}，θ_{10}）。因此一个全局最小值从 Cov（ε（t，θ_1，θ_2））中得到，即使 θ_2 为任何固定值，它也可能与 θ_{20} 完全不同。直接应用这个属性设置 G_l（z^{-1}，θ_2）$=1$，实际上一个参考模型和过程模型的一致性是可以实现的。

4.2.4　PEM 渐近方差属性

一旦对系统模型的参数进行估计，确定估计的参数的不确定性就很必要。在线性回归的情况下，参数估计的协方差可以简单地确定为如前面所讨论 σ^2（$\Phi^{\mathrm{T}}\Phi$）$^{-1}$ 在涉及非线性优化的 PEM 中，协方差是通过预测误差的梯度来确定的。推导的细节参考 Söderström 和 Stoica（1989）的研究，结果概要如下。

假设

1）$\{u_t,\ y_t\}$，$t=1,\ 2,\ \cdots,\ N$ 这些数据是平稳过程。

2）输入为持续激励。

3）Hessian J''（θ）是非奇异（非单数）。在周围 J（θ）其中，J（θ）的损失函数最小点局部最小。

4）G_p（z^{-1}；θ）G_l（z^{-1}；θ）是参数矢量 $\boldsymbol{\theta}$ 的微分函数。

5）该系统是可辨识的。

然后，将得到 PEM 的以下属性：

- $\sum_\infty = \lim_{N\to\infty}\sum_N(\theta) = E\varepsilon(t,\theta)\varepsilon^{\mathrm{T}}\varepsilon(t,\theta)$。
- 渐近参数估计 $\hat{\theta}_N$ 遵循正态分布：

$$\sqrt{N}(\hat{\theta}_N-\theta_0)\sim N(0,P) \tag{4.81}$$

式中

$$P=\lambda^2[E\psi(t,\theta_0)\psi^{\mathrm{T}}(t,\theta_0)]^{-1}$$

式中

$$\psi_{(}t,\theta) = -\left(\frac{\partial\varepsilon(t,\theta)}{\partial\theta}\right)^{\mathrm{T}}$$

P 的近似估计可为

$$\hat{P}=\lambda^2\left[\frac{1}{N}\sum_{t=1}^{N}\psi(t,\hat{\theta}_N)\psi^{\mathrm{T}}(t,\hat{\theta}_N)\right]^{-1}$$

作为例子，将证明，作为 PEM 的特殊情况，估计 ARX 模型参数的协方差，根

据 PEM 的渐近方差表达式派生是相同的，直接从线性回归推导。

考虑表示回归方程的形式 ARX 模型：

$$y_t = \varphi^T(t)\theta + e_t$$

由给出的从 $t=1\sim N$ 的数据可估计参数 θ 的协方差计算如下：

$$\mathrm{Cov}(\hat{\theta}) = \lambda^2(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1} \tag{4.82}$$

式中

$$\boldsymbol{\Phi} = \begin{pmatrix} \varphi^{\mathrm{T}}(1) \\ \vdots \\ \varphi^{\mathrm{T}}(N) \end{pmatrix}$$

结合这两个方程得出：

$$\mathrm{Cov}(\hat{\theta}) = \lambda^2\left[\sum_{t=1}^{N}\varphi(t)\varphi^{\mathrm{T}}(t)\right]^{-1} \tag{4.83}$$

现在，我们应用 PEM 渐近协方差估计最小二乘估计的 ARX 模型和直接来自最小二乘回归验证的结果是相同的。应用 PEM 的渐近协方差到 ARX 模型，最优预测首先推导为

$$\hat{y}(t|t-1) = \varphi^{\mathrm{T}}(t)\theta$$

这被认为是最佳预测，因为如果模型是最优的它会提供白噪声作为其预测误差。现在，预测误差可表示为

$$\varepsilon(t,\theta) = y_t - \hat{y}_t = y_t - \varphi^{\mathrm{T}}(t)\theta$$

结果是

$$\psi(t,\theta) = -\left(\frac{\partial\varepsilon(t,\theta)}{\partial\theta}\right)^{\mathrm{T}} = \varphi(t)$$

因此

$$\hat{P} = \lambda^2\left[\frac{1}{N}\sum_{t=1}^{N}\psi(t,\hat{\theta}_N)\psi^{\mathrm{T}}(t,\hat{\theta}_N)\right]^{-1} = \lambda^2\left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi(t)\right]^{-1}$$

当 $N\to\infty$ 时，$\hat{P}=P$。根据式（4.81）有

$$\mathrm{Cov}(\hat{\theta}) = \frac{1}{N}P = \frac{1}{N}\lambda^2\left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi(t)\right]^{-1} = \lambda^2\left[\sum_{t=1}^{N}\varphi(t)\varphi(t)\right]^{-1}$$

这同式（4.83）所示的线性回归得到的结果是一样的。

4.2.5　非线性辨识

这个类别中的最后一部分内容介绍了非线性模型结构的范围。本节将专注 PEM 应用于特定的结构。这是非线性 ARX 模型或 NARX，其中噪声是附加白噪声。回想 NARX 模型：

$$y_t = f(y_{t-1},\cdots,y_{t-p},u_t,\cdots,u_{t-q}) + e_t$$

式中，$f(y_{t-1},\cdots,y_{t-p},u_t,\cdots,u_{t-q})$ 没有特殊的结构，除非非线性模型来源于第一性原理。

$f(y_{t-1}, \cdots, y_{t-p}, u_t, \cdots, u_{t-q})$ 通常是未知的，不能写出它的一般表达式。绕过这个问题的一个常用的方法是使用基函数，通过级数展开来近似 $f(y_{t-1}, \cdots, y_{t-p}, u_t, \cdots, u_{t-q})$。

4.2.5.1　基函数

定义：

$$Z^t = \{y_t, y_{t-1}, \cdots, u_t, u_{t-1}, \cdots\}$$

类似于 PEM 的线性系统，需要找到最佳的一步超前预测：

$$\hat{y}(t|t-1) = f(Z^{t-1}, \theta)$$

类似于线性回归，定义一个回归，使得

$$\varphi(t) = \varphi(Z^{t-1})$$

然后，预测为

$$f(Z^{t-1}, \theta) = f(\varphi(t), \theta)$$

由于没有一般形式的非线性函数 $f(\varphi(t), \theta)$，很自然地通过级数展开以参数化非线性函数：

$$f(\varphi(t), \theta) = \sum_{k=1}^{n} \alpha_k f_k(\varphi(t)), \quad \theta = [\alpha_1 \cdots \alpha_n]^{\mathrm{T}}$$

式中，f_k 被称为一个基函数。

f_k 一般是从一个“母基函数”得到的，记为 $\kappa(x)$：

$$f_k(\varphi(t)) = f_k(\varphi(t), \beta_k, \gamma_k) = \kappa(\beta_k(\varphi(t) - \gamma_k))$$

有些基函数示例如下：

- 泰勒级数 $\kappa(x) = x^k$。
- 傅里叶级数 $\kappa(x) = \cos(x)$。
- 分段常数函数：

$$\kappa(x) = \begin{cases} 1 & 0 \leqslant x < 1 \\ 0 & \text{其他} \end{cases}$$

或分段常数函数的平滑版本，即钟形（高斯）函数：

$$\kappa(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$$

- 阶跃函数：

$$\kappa(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geqslant 0 \end{cases}$$

或阶梯函数的平滑版本，即（S 形）函数：

$$\kappa(x) = \sigma(x) = \frac{1}{1 + \mathrm{e}^{-x}}$$

这些母基函数可以近似地认为是结构材料的“砖”。一个非线性函数可以类似地认为这是从砖建造的建筑物。各种形状和尺寸的砖块将必须被移动到施工现场和

被组装以形成一个建筑物。这也适用于用母基函数表达非线性函数。

4.2.5.2　非线性预测构造

要表示的非线性函数，必须用适当的形状和比例的基础功能进行构造，然后重新定位。这通过以下结构式完成：

$$f(\varphi(t),\theta) = \sum_{k=1}^{n}\alpha_k\kappa(\beta_k(\varphi(t)-\gamma_k))$$

现在 θ 的参数化的特征在于 3 种类型的参数：

- 坐标，α_k；
- 缩放参数，β_k；
- 位置参数，γ_k。

其中坐标参数 α_k 分配权重贡献给每个单独的基函数的非线性函数 f（φ（t），θ）。

现在，问题是，能否在基函数级数逼近真实的非线性函数 f（φ（t），θ）。“对于 κ（x）几乎任何选择除了对一个多项式，扩展可以逼近任何合理的函数 g_0（φ（t））以及对足够大的 n”（Ljung，1999）。然而，潜在的问题是非线性函数的过度拟合。由于基函数的数量增加，因此没有参数确定。系统辨识中的简约性原则，后面将简短介绍，决定了该模型的预测性将会恶化。

4.2.5.3　多层网络

在非线性系统辨识中，通常通过以捕捉复杂的非线性增加额外层来扩大基础功能。基函数的输出是

$$\varphi_k^{(2)}(t)=f_k(\varphi(t))=\kappa(\varphi(t),\beta_k,\gamma_k)$$

将它们集中到一个矢量中：

$$\boldsymbol{\varphi}_t^{(2)}(t)=[\varphi_1^{(2)}(t),\varphi_2^{(2)}(t),\cdots,\varphi_n^{(2)}(t)]^{\mathrm{T}}$$

而非采取 $\varphi_k^{(2)}$（t）的线性组合作为预测，可以把它们作为新的回归，并把它们插入到基函数的另一层：

$$f(\varphi(t),\theta) = \sum_{l}\alpha_l^{(2)}\kappa(\varphi^{(2)}(t),\beta_l^{(2)},\gamma_l^{(2)})$$

在神经网络的文献中，上述结构近似的预测被称为双隐层网络。同样，人们可以获得更多的层。虽然包括额外的层将提供更好的拟合数据，但它通常会增加过度参数化问题的可能性。

4.3　模型验证

当遇到系统辨识问题时，基础模型结构的先验知识可能是有限的。当从数据中搜索“正确”的模型结构时，人们可能不得不面对以下问题：哪些候选模型结构应该选择？选择的模型是否过于复杂？

4.3.1　模型结构选择

对于线性系统识别，各模型的结构都可以考虑，例如 ARX、ARMAX、OE、BJ、

PE 和状态空间模型。对于相同的模型结构，有不同的模型命令。一些统计测试可以用来确定基于预测误差 $\varepsilon(t,\hat{\theta}_N)$ 适当的模型结构或顺序。参数估计 $\hat{\theta}_N$ 计算预测误差称为残差。回想一下式（4.56）和式（4.57）及最优预测误差表达式：

$$
\begin{aligned}
\varepsilon(t,\theta) &= \Phi_u(z^{-1};\theta,\theta_0)u_t + (\Phi_e(z^{-1};\theta,\theta_0) - I)e_t + e_t \\
\Phi_u(z^{-1};\theta,\theta_0) &= G_l^{-1}(z^{-1};\theta)[G_p(z^{-1};\theta_0) - G_p(z^{-1};\theta)] \\
\Phi_e(z^{-1};\theta,\theta_0) &= G_l^{-1}(z^{-1};\theta)G_l(z^{-1};\theta_0)
\end{aligned}
\tag{4.84}
$$

基于上面的方程可以看成两个测试：

1. 自相关测试

从式（4.84）中可知，如果估计模型等于真实值，前两个方面的预测误差将变为零。因此，如果两个过程和干扰模型是完美的，那么预测误差必须是白噪声。在实践中，不会拒绝一个信号是白噪声的假设，如果除了在滞后 0 超过 99% 的自相关点落入$\left[-\frac{3}{\sqrt{N}}, \frac{3}{\sqrt{N}}\right]$范围内。自相关是归一化的自相关，它可以计算如下：

$$
\rho_\varepsilon(\tau) = \frac{r_\varepsilon(\tau)}{r_\varepsilon(0)} \tag{4.85}
$$

式中 $r_\varepsilon(\tau)$ 代表与 τ 滞后的预测误差自协方差。如果超过 1% 的 $\rho_\varepsilon(\tau)$ 满足 $|\rho_\varepsilon(\tau)| > \frac{3}{\sqrt{N}}$，预测误差是不被认为是白噪声并且估计的模型是无效的。

2. 互相关测试

由式（4.84）可知，如果该进程模型估计是完美的，$\Phi_u(q^{-1},\theta)$ 将成为零，而且会预测误差和输入之间没有交叉相关。互相关是归一化的互协方差，其可以被计算为

$$
\rho_{\varepsilon u}(\tau) = \frac{r_{\varepsilon u}(\tau)}{\sqrt{r_\varepsilon(0)r_u(0)}} \tag{4.86}
$$

式中，$r_{\varepsilon u}(\tau)$ 表示预测误差和输入之间的互协方差与 τ 滞后。$\rho_{\varepsilon u}(\tau)$ 满足 $|\rho_{\varepsilon u}(\tau)| > \frac{3}{\sqrt{N}}$时，估计过程模型是无效的。

作为一个经验法则，互相关测试用于过程模型失效，而自相关测试用于无效的进程和扰动模型。

如果估计的模型失效，选择不同的模型结构或顺序，并且重复上述过程。在实践中，这两项测试都必须对所有模型结构进行除 OE 模型结构，因为自相关测试并不适用于 OE 模型的失效。仅仅互相关测试可以应用到 OE 模型。

4.3.2　简约原则

简约原则说的是两个或两个以上相互竞争的模型，这一切能很好地解释数据，该模型具有最小数量的独立参数选择。这样的选择会带给预测方面最佳精度。

假定 $\hat{\boldsymbol{\theta}}_N$ 表示使用 PEM 从过去的数据估计出的参数矢量。进一步假设该系统属于所考虑的模型结构。这意味着存在一个真实参数矢量 $\boldsymbol{\theta}_0$。

一个适当的标量量度可用于评估该模型与 $\hat{\boldsymbol{\theta}}_N$ 相关联的优度。例如，这项测量可以是

$$W(\hat{\theta}_N) = E\varepsilon^{\mathrm{T}}(t,\hat{\theta}_N)\varepsilon(t,\hat{\theta}_N)$$

其中预测误差基于未来数据，并且期望是关于从过去的数据估计的参数的未来噪声的条件。

然后，其表示如下（Söderström 和 Stoica，1989）：

$$EW(\hat{\theta}_N) \approx \lambda^2(1+p/N)$$

式中，p 是 $\hat{\theta}_N$ 的维数；N 是用于估计模型和 $\lambda^2 = Ee_t^2$ 数据点的数量。

期望是相对于过去的数据或 $\hat{\theta}_N$ 的分布，这是从过去的数据估计得到的。这个结果表示，p/N 的相对量随预期误差方差增大。因此，存在于使用不利模型不必要的许多参数。这可以被看作简约原则的声明。

4.3.3　模型结构比较

4.3.3.1　χ^2 测试

让 μ_1 和 μ_2 为两个模型结构，这样，$\mu_1 \subset \mu_2$。此外，让损失函数 $J^{(i)}$ 表示结构中 μ_i（$i=1$，2）的最小 J（$\hat{\theta}_N$），并让 μ_i 有参数 p_i。损失函数已经被定义为

$$J(\hat{\theta}_N) = \frac{1}{N}\varepsilon^{\mathrm{T}}(t,\hat{\theta}_N)\varepsilon(t,\hat{\theta}_N)$$

如果 μ_1 足够大以包括在真实系统中，那么

$$x = N\frac{J^{(1)} - J^{(2)}}{J^{(2)}} \sim \chi^2(p_2 - p_1)$$

这个结果可以用于模型结构选择的统计测试。在显著性水平 α 中，较小的模型结构 μ_1 被选中，如果

$$x \leqslant \chi_\alpha^2(p_2 - p_1)$$

式中，χ_α^2 为 α 的 χ^2 分布的显著水平。

4.3.3.2　AIC

赤池信息标准（Akaike's Information Criterion，AIC）由下式给出：

$$\mathrm{AIC} = N\log J(\hat{\theta}_N) + 2p$$

式中，$N\log J$（$\hat{\theta}_N$）是模型参数（p）的数量逐渐增加而普遍下降；$2p$ 随 p 的数量增加线性增加。

例如在某些方面标准处罚损失函数 J（$\hat{\theta}_N$）随着模型尺寸的增加而减小。该模型的结构赋予这一标准的最小值被选中。

4.3.3.3　FPE 准则

最终预报误差（FPE）准则由下式给出：

$$\mathrm{FPE} = J(\hat{\theta}_N)\frac{1+p/N}{1-p/N}$$

如下面的内容中所展示的，FPE 可以来自简约性原则：

$$EW(\hat{\theta}_N) \approx \lambda^2(1+p/N) \tag{4.87}$$

按照单输出系统，λ^2 的估计：

$$\begin{aligned}\hat{\lambda}^2 &= \frac{1}{N-p}\sum_{t=1}^{N}\varepsilon^2(t,\hat{\theta}_N)\\ &= \frac{N}{N-p}\frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\hat{\theta}_N)\\ &= \frac{1}{1-p/N}J(\hat{\theta}_N)\end{aligned}$$

代换式（4.87）的 $\hat{\lambda}^2$ 可得

$$EW(\hat{\theta}_N) \approx \lambda^2(1+p/N) \approx J(\hat{\theta}_N)\frac{1+p/N}{1-p/N}$$

这就是所谓的 FPE。

4.4　经验法

知识系统辨识的过程和经验对于一个成功的系统识辨程序是必不可少的，应始终考虑经验。例如，知道该过程是否具有非线性、随时间变化的特性、漂流紊乱等，可以在系统辨识过程起到重要的作用。以下是一些可以被认为是进行系统辨识时的实际问题。

4.4.1　非零处理

由于在系统识别中使用模型结构，所有变量的性质鉴定应该具有零均值。这可以很容易地理解，因为在模型中，式（4.38）没有使用常数项，计算非零均值，白噪声项 e_t 等被假定为具有零均值。然而，如果不减去输入和输出数据的均值，则所识别的模型将比通过一个因数（$1-z^{-1}$）的实际模型的值要高。例如，考虑一个真实的模型，可以产生非零均值数据：

$$A_1(z^{-1})y_t = B_1(a^{-1})u_t + C_1(z^{-1})e_t + m$$

通过辨识算法承认该模型只采用以下形式：

$$A(z^{-1})y_t = B(a^{-1})u_t + C(z^{-1})e_t$$

通过乘以真实模型（$1-z^{-1}$）得出：

$$(1-z^{-1})A_1(z^{-1})y(t) = (1-z^{-1})B_1(a^{-1})u_t + (1-z^{-1})C_1(z^{-1})e_t$$

常数项消失，所得到的模型具有一个辨识算法中使用的形式。然而，模型阶数

已增加。在实践中，不去除均值会有不利的影响。

4.4.2 漂移干扰处理

漂移的干扰往往体现在输出数据 y_t 为平均值的漂移。漂移干扰可能由随机游走干扰，这是建模实践中最常见的干扰：

$$v_t = \frac{H(z^{-1})}{1 - z^{-1}} e_t$$

模型结构应为

$$y_t = G(z^{-1}) u_t + \frac{H(z^{-1})}{1 - z^{-1}} e_t$$

或者

$$\Delta y_t = G(z^{-1}) \Delta u_t + H(z^{-1}) e_t$$

式中，$\Delta = (1 - z^{-1})$，也就是说，如果该数据具有漂移均值，人们可以在应用辨识算法之前运用差分运算。然而，差分数据是相当于施加极端的高通滤子，可以显著增加噪声对识别的影响。

4.4.3 鲁棒性

当记录实验数据时，可能偶尔会出现大的测量误差。这种错误可以通过干扰、传感器故障等被引入。所得到的异常数据点被称为离群值。如果不采取具体行动，离群值将大大影响估计模型。异常值趋向于出现在序列中的尖峰预测误差，并会因此对损失函数有很大的贡献。这是很难直接从动态数据中检测离群值的。事实上，如果输出 y_t 被视为离群值是相当误导的，作为尖峰可能已经引起在输入方面的根本变化。异常值应该用残差来确定，它始终是绘制残差来检查是否有从零到任何大型漂移残差的一个好办法。如果被检测到异常，它不应该被简单地删除，因为删除可能会改变数据的动态信息。通常，异常值被替换为它的预测输出，但是这样做也将面临一个“抓22（catch 22）”的局面。计算残差或预测误差，需要有模型。有异常值的存在，从原始数据中确定的模型不会好。解决的办法是采用迭代过程进行异常检测和模型建立。

4.4.4 额外的模型验证

模型验证是关注所获得的模型是否足够。另外如前面所讨论的模型验证，它也是重要的检查先验假设。

• 线性度的测试：如果可能的话，实验应重复输入信号的另一个幅度，以验证对哪些操作范围内的线性模型是足够的。作为一个简单的测试，用户应尽量兼顾正的和负的步骤来检查响应的对称或步骤输入的大小不同，以检查是否响应正比于输入的幅度的变化。

• 时不变的测试：这是与交叉验证非常相似的测试。一个典型的数据被分离成两部分，适合模型中的每个部分的数据，然后比较这两个模型。如果这个过程是时不变的，那么这两个模型应该是相似的。另外，也可以适用于一些参数变化检测算法来验证模型参数的时变。

• 测试反馈的存在：如果在数据中没有反馈，模型拟合数据良好，那么残差应该是过去和未来的输入独立。然而，如果有一个反馈，那么残差将与未来的输入相关，就是说对于 $\tau \leqslant 0$ $r_{\varepsilon u}(\tau)=E\varepsilon_t u_{t-\tau} \neq 0$。

4.5　闭环辨识

闭环辨识的主要挑战是干扰和输入之间相关性的存在。由 θ_0 来替换图 4.1 中的 θ，就可以得到下面两个方程来描述的闭环响应：

$$y_t = G_p(z^{-1},\theta_0)u_t + G_l(z^{-1},\theta_0)e_t$$

$$u_t = -G_c(z^{-1})y_t + G_c(z^{-1})r_t$$

这些方程可以变换为

$$y_t = G_p(z^{-1},\theta_0)u_t + G_l(z^{-1},\theta_0)e_t \tag{4.88}$$

$$y_t = -G_c^{-1}(z^{-1})u_t + r_t \tag{4.89}$$

给定一组输入和输出数据 u_1，y_1，…，u_N，y_N，通过 PEM 估算出模型的参数最小的预测误差。回答由式（4.88）和式（4.89）给出的扰动项哪两个模型拟合数据较好，则应是两个方程的右边第二项。很明显，如果 $r_t=0$（没有外部激励），所选模型结构与 $G_c^{-1}(z^{-1})$ 是相同的，式（4.89）将完全适合数据。这种情况导致了辨识过程模型 G_p（z^{-1}，θ_0）的彻底失败。另一方面，如果 $e_t=0$ 或者 r_t 的大小比 $G_l(z^{-1}，\theta_0)e_t$ 中的大很多，那么它的大小就可以被测量，例如，通过方差，那么式（4.88）将更好地拟合数据。其结果是，PEM 会收敛到由式（4.88）给出的模型。后一种情况下，使得过程模型 $G_p(z^{-1}，\theta_0)$辨识。

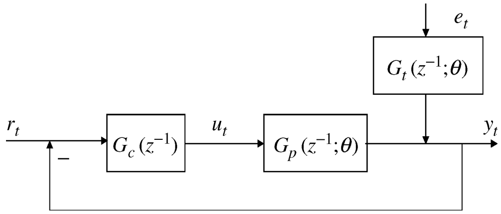


图 4.1　闭环系统

除了前面讨论过的两个明显的情况，还有其他的情况可以对过程进行辨识。例如，在无 r_t 情况下，如果模型结构选择的与 $G_c^{-1}(z^{-1})$足够不同和与 $G_p(z^{-1}，\theta_0)$ 足够接近，接下来，通过模型不匹配引入的“干扰”可能会取代由 $G_l(z^{-1}，\theta_0)e_t$ 造成的干扰。其结果是，式（4.88）将更好地适合数据。有几种方法来产生这种情况。例如，对于 $G_c(z^{-1})$的模型阶数比 $G_p(z^{-1}，\theta_0)$的更高，而用于辨识的模型阶数被选择对于 $G_p(z^{-1}，\theta_0)$是相同的。另外，$G_c(z^{-1})$是非线性的或时变的，而 $G_p(z^{-1}，\theta_0)$是线性的或时不变的，并且模型结构辨识选择为线性的或时不变的。

前面讨论的情况可能使得在没有外部激励 r_t 的情况下 $G_p(z^{-1},\ \theta_0)$ 辨识。然而，最可靠的条件下仍然会存在外部激励。以下内容侧重于讨论后者的条件。

4.5.1　直接闭环辨识

当输入和输出数据的 u_t 和 y_t 从闭环系统中采样被直接用于辨识时，闭环辨识问题被视为一个“开环”辨识问题。辨识采用闭环数据这种方法被称为直接闭环辨识法或简称直接辨识法。如果 $G_p(z^{-1},\ \theta)$ 中有至少一个样本的时间延迟，并且 r_t 是持续激励的全阶，则当 $N\rightarrow\infty$ 时，$G_p(z^{-1},\ \theta)$ 和 $G_l(z^{-1},\ \theta)$ 将直接应用到 PEM 的输入和输出数据 u_t、y_t 收敛到其真实值。

为了便于说明，考虑一个单输入单输出系统，其中真正的过程模型和干扰模型分别是 $G_p(z^{-1},\ \theta)$ 和 $G_l(z^{-1},\ \theta)$。根据图 4.1 所示，输入可表示为

$$u_t=\frac{G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}r_t+\frac{-G_l(z^{-1},\theta_0)G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}e_t \tag{4.90}$$

将式（4.88）中描述的真实系统代入到式（4.43）中描述的预测误差得出：

$$\begin{aligned}\varepsilon(t,\theta)&=\frac{1}{G_l(z^{-1};\theta)}[y_t-G_p(z^{-1};\theta)u_t]\\&=\frac{1}{G_l(z^{-1};\theta)}[G_p(z^{-1};\theta_0)-G_p(z^{-1};\theta_0)]u_t+\frac{G_l(z^{-1};\theta_0)}{G_l(z^{-1};\theta)}e_t\end{aligned}$$

的确，因为采用直接鉴定法，这与最优预测和预测误差的表达式计算开环辨识相同。将上面方程中的结果代入式（4.90）得到

$$\begin{aligned}\varepsilon(t,\theta)=&\frac{1}{G_l(z^{-1};\theta)}[G_p(z^{-1};\theta_0)-G_p(z^{-1};\theta)]\Big[\frac{G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}r_t\\&+\frac{-G_l(z^{-1},\theta_0)G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}e_t\Big]+\frac{G_l(z^{-1};\theta_0)}{G_l(z^{-1};\theta)}e_t\end{aligned} \tag{4.91}$$

定义灵敏度函数：

$$S(z^{-1},\theta)\triangleq\frac{1}{1+G_c(z^{-1})G_p(z^{-1},\theta)}$$

$$S(z^{-1},\theta_0)\triangleq\frac{1}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}$$

式（4.91）可以简化为

$$\begin{aligned}\varepsilon(t,\theta)=&\frac{1}{G_l(z^{-1};\theta)}[G_p(z^{-1};\theta_0)-G_p(z^{-1};\theta)]S(z^{-1},\theta_0)G_c(z^{-1})r_t\\&+\frac{S(z^{-1},\theta_0)}{S(z^{-1},\theta)}\frac{G_l(z^{-1};\theta_0)}{G_l(z^{-1};\theta)}e_t\end{aligned} \tag{4.92}$$

如下恒等式：

$$G_p(z^{-1}=0,\theta)=G_p(z^{-1}=0,\theta_0)=0$$

意味着

$$S(z^{-1}=0,\theta)=S(z^{-1}=0,\theta_0)=1$$

回顾

$$G_l(z^{-1}=0,\theta)=G_l(z^{-1}=0,\theta_0)=1$$

扩展式（4.92）的右边的第二项脉冲形式将得到

$$\begin{aligned}\frac{S(z^{-1},\theta_0)}{S(z^{-1},\theta)}\frac{G_l(z^{-1};\theta_0)}{G_l(z^{-1};\theta)}e_t &=(1+\text{与 } z^{-1},z^{-2},\cdots\text{相关项})e_t\\ &=e_t+\text{与 } e_{t-1},e_{t-2},\cdots\text{相关项}\end{aligned}$$

由于 r_t 和 e_t 是独立的，e_t 是白噪声，运用方差作用式（4.92）得出

$$\begin{aligned}\mathrm{Var}(\varepsilon(t,\theta))=&\mathrm{Var}\left(\frac{1}{G_l(z^{-1};\theta)}[G_p(z^{-1};\theta_0)-G_p(z^{-1};\theta)]S(z^{-1},\theta_0)G_c(z^{-1})r_t\right)\\ &+\mathrm{Var}(e_t+\text{与 } e_{t-1},e_{t-2}\cdots\text{相关项})\\ \geqslant&\mathrm{Var}(e_t)\end{aligned}$$

这表明，Var（e_t）是 Var（ε（t，θ））的最小值。

显然，人们可以从式（4.92）看出，如果 $G_p(z^{-1},\theta)=G_p(z^{-1},\theta_0)$ 并且 $G_l(z^{-1},\theta)=G_l(z^{-1},\theta_0)$，那么

$$\mathrm{Var}(\varepsilon(t,\theta))=\mathrm{Var}(e_t)$$

也就是说，得到 Var(ε（t，θ)) 下界。另一方面，如果得到了下界，然后式（4.92）右侧的第一项必须为零，这意味着 $G_p(z^{-1};\theta)=G_p(z^{-1};\theta_0)$，因此，$S(z^{-1};\theta)=S(z^{-1};\theta_0)$。结果是，必须有 $G_l(z^{-1};\theta)=G_l(z^{-1};\theta_0)$，使 Var（$\varepsilon$（$t$，$\theta$））= Var($e_t$)。因此，得到 Var（$\varepsilon$（$t$，$\theta$））的下界对于辨识的一致性是一个充分必要条件。

从上面讨论的推导可以得出结论，对于直接的一致性闭环识别相当于在进程和扰动模型的同时一致性。如果它们中的一个不能达到一致，另外一个可能也失败。例如，在式（4.92）中，如果 $G_l(z^{-1};\theta)\neq G_l(z^{-1};\theta_0)$，则对于最小化的 ε（t，θ），$S(z^{-1},\theta)$ 不一定等于 $S(z^{-1},\theta_0)$。因此，$G_p(z^{-1};\theta)$ 不一定等于 $G_p(z^{-1};\theta_0)$。出于这个原因，当选择模型结构直接闭环辨识时，必须确保这两个过程和干扰模型的结构应该是相同或跟一般比真实的。实际上，这意味着，对于线性模型，如果实际模型结构并不之前已知，只有 PEM 和 BJ 模型结构可考虑。

4.5.2　间接闭环辨识

最直观解决闭环识别问题的方法是首先确定闭环模式，并从该过程模型提取。这个方法也被称为间接闭环辨识方法。

根据图 4.1，闭环响应设定点由下式给出：

$$\begin{aligned}y_t&=\frac{G_c(z^{-1})G_p(z^{-1},\theta)}{1+G_c(z^{-1})G_p(z^{-1},\theta)}r_t+\frac{G_l(z^{-1},\theta)}{1+G_c(z^{-1})G_p(z^{-1},\theta)}e_t\\ &\triangleq M(z^{-1},\theta)r_t+N(z^{-1},\theta)e_t\end{aligned}$$

使用设定点 r_t 和输出数据 y_t 辨识是一个开环的识别问题，r_t 和 e_t 是独立的。那么 $M(z^{-1},\theta)$ 和 $N(z^{-1},\theta)$ 可以被首先确定，然后 $G_p(z^{-1},\theta)$ 和 $G_l(z^{-1},\theta)$ 可以分别从 $M(z^{-1},\theta)$ 和 $N(z^{-1},\theta)$ 中获得即

$$G_p(z^{-1},\theta)=\frac{1}{\dfrac{G_c(z^{-1})}{M(z^{-1},\theta)}-G_c(z^{-1})}$$

$$G_l(z^{-1},\theta)=N(z^{-1},\theta)(1+G_c(z^{-1})G_p(z^{-1},\theta))$$

从这种方法估计的 $G_p(z^{-1},\theta)$ 和 $G_l(z^{-1},\theta)$ 可能是相对高的阶。进行进一步的模型降阶是必要的。

例 4.5.1　考虑如图 4.1 所示的闭环系统，其中真实过程和干扰如下式所示：

$$y_t=\frac{b_0z^{-2}}{1+a_0z^{-1}}u_t+\frac{1+c_0z^{-1}}{1+d_0z^{-1}}e_t \tag{4.93}$$

控制器的传递函数由下式给出：

$$G_c(z^{-1})=\frac{k}{1-z^{-1}}$$

式中，r_t 是设定点足够阶数的持续激励。闭环辨识方法应用于具有相同模型的结构，为那些真正的过程和干扰。

1）通过应用直接 PEM，什么将是对提前一步预测误差的方差下界？如果这个下界是可实现的，那么估计过程模型和干扰模型会是它们的真实值吗？

2）现在假设 $r_t=e_t+\delta_t$，其中 δ_t 也是一个白噪声序列且 e_t 是独立的。如果间接辨识方法应用于一个适当的预测误差的方法，过程和干扰模型（如 G_p 和 G_l）的估计会是一致的吗？

回想一下，一个线性预测器的形式为

$$y_t=L_1(z^{-1},\theta)y_t+L_2(z^{-1},\theta)u_t \tag{4.94}$$

注意，$G_p(z^{-1},\theta)$ 和 $G_l(z^{-1},\theta)$ 模型是在预测可以推导的基础上，但 $G_p(z^{-1},\theta)$ 和 $G_l(z^{-1},\theta)$ 是真实的，但是未知模型。同理，在开环的情况下，最优的预测是

$$L_1(z^{-1},\theta)=I-G_l^{-1}(z^{-1},\theta)=\left[1-\frac{1+dz^{-1}}{1+cz^{-1}}\right] \tag{4.95}$$

$$L_2(z^{-1},\theta)=G_l^{-1}(z^{-1},\theta)G_p(z^{-1},\theta)=\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{bz^{-1}}{1+az^{-1}}\right] \tag{4.96}$$

因此，该预测误差由下式给出：

$$\begin{aligned}\varepsilon(t,\theta)&=y_t-\hat{y}(t\,|\,t-1)\\&=y_t-(I-G_l^{-1}(z^{-1},\theta))y_t-G_l^{-1}(z^{-1},\theta)G_p(z^{-1},\theta)u_t\\&=\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]y_t-\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{bz^{-1}}{1+az^{-1}}\right]u_t\end{aligned}$$

$$= \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{b_0 z^{-2}}{1+a_0 z^{-1}}u_t + \frac{1+c_0 z^{-1}}{1+d_0 z^{-1}}e_t\right] - \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{bz^{-1}}{1+az^{-2}}\right]u_t$$

这使

$$\varepsilon(t,\theta) = \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{b_0}{1+a_0 z^{-1}} - \frac{bz^{-1}}{1+az^{-1}}z^{-2}\right]u_t + \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{1+c_0 z^{-1}}{1+d_0 z^{-1}}\right]e_t \quad (4.97)$$

这个预测误差表现在 u_t 上，它需要被变换成实际的外生扰动，r_t 与 e_t 是独立的。

用 G_p（z^{-1}，θ）和 G_l（z^{-1}，θ）的真实值 G_p（z^{-1}，θ_0）和 G_l（z^{-1}，θ_0）分别来代替，输入 u_t 可以从图 4.1 这样得到：

$$u_t = \frac{G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}r_t + \frac{-G_l(z^{-1},\theta_0)G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1},\theta_0)}e_t$$
$$= S(z^{-1},\theta_0)G_c(z^{-1})r_t - S(z^{-1},\theta_0)G_c(z^{-1})G_l(z^{-1},\theta_0)e_t \quad (4.98)$$

式（4.97）代入式（4.98）得到

$$\varepsilon(t,\theta) = \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{b_0}{1+a_0 z^{-1}} - \frac{bz^{-1}}{1+az^{-1}}\right]S(z^{-1},\theta_0)G_c(z^{-1})r_t + \left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{1+c_0 z^{-1}}{1+d_0 z^{-1}}\right]\frac{S(z^{-1},\theta_0)}{S(z^{-1},\theta)}e_t \quad (4.99)$$

通过脉冲响应的形式写式（4.99），可以看到下面的表达式：

$$\varepsilon(t,\theta) = e_t + \text{与 } e_{t-1}, e_{t-2}, \cdots \text{相关项} + \text{与 } r_t \text{ 相关项}$$

其结果是

$$\mathrm{Var}(\varepsilon(t,\theta)) \geqslant \mathrm{Var}(e_t)$$

r_t 和 e_t 是独立的，闭环的预测误差方差表达式从式（4.99）推导为

$$\mathrm{Var}(\varepsilon(t,\theta)) = \mathrm{Var}\left(\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{b_0}{1+a_0 z^{-1}} - \frac{bz^{-1}}{1+az^{-1}}\right]S(z^{-1},\theta_0)G_c(z^{-1})r_t\right) + \mathrm{Var}\left(\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{1+c_0 z^{-1}}{1+d_0 z^{-1}}\right]\frac{S(z^{-1},\theta_0)}{S(z^{-1},\theta)}e_t\right) \quad (4.100)$$

可以看出，通过让 $b = b_0$ 和 $a = a_0$，式（4.100）的第一项将是零，因此

$$\mathrm{Var}(\varepsilon(t,\theta)) = \mathrm{Var}\left(\left[\frac{1+dz^{-1}}{1+cz^{-1}}\right]\left[\frac{1+c_0 z^{-1}}{1+d_0 z^{-1}}\right]\frac{S(z^{-1},\theta_0)}{S(z^{-1},\theta)}e_t\right) \quad (4.101)$$

如果进一步令 $d = d_0$ 和 $c = c_0$，则式（4.101）简化为

$$\mathrm{Var}(\varepsilon(t,\theta)) = \mathrm{Var}(e_t)$$

这是 $\mathrm{Var}(\varepsilon$（T，θ）) 的下界。它也可以被类似地证实当下界是可实现时，的确 $G_p(z^{-1}$，$\theta) = G_p$（z^{-1}，θ_0）并且 $G_l(z^{-1}$，$\theta) = G_l$（z^{-1}，θ_0）。

现在，考虑间接闭环辨识方法 $r_t = e_t + \delta_t$，其中 δ_t 也是白噪声且 e_t 是独立的项。对于闭环形式，用真实的过程和干扰模型，根据 r_t 表达 y_t：

$$y_t=\frac{G_p(z^{-1},\theta_0)G_c(z^{-1})}{1+G_p(z^{-1},\theta_0)G_c(z^{-1})}r_t+\frac{G_l(z^{-1},\theta_0)}{1+G_p(z^{-1},\theta_0)G_c(z^{-1})}e_t$$
$$\triangleq M(z^{-1},\theta_0)r_t+N(z^{-1},\theta_0)e_t \tag{4.102}$$

表达 y_t 关于 r_t 与假定过程和扰动模型的闭环形式：

$$y_t=\frac{G_p(z^{-1},\theta)G_c(z^{-1})}{1+G_p(z^{-1},\theta)G_c(z^{-1})}r_t+\frac{G_l(z^{-1},\theta)}{1+G_p(z^{-1},\theta)G_c(z^{-1})}e_t$$
$$\triangleq M(z^{-1},\theta)r_t+N(z^{-1},\theta)e_t \tag{4.103}$$

单步向前最佳预测是基于假设的模型设计的，并且与开环系统是相同的（除了它现在 r_t 中表示，因此是间接的）：

$$\hat{y}(t|t-1)=(I-N^{-1}(z^{-1},\theta))y_t+N^{-1}(z^{-1},\theta)M(z^{-1},\theta)r_t \tag{4.104}$$

预测误差是

$$\begin{aligned}\varepsilon(t,\theta)&=y_t-\hat{y}(t|t-1)\\&=y_t-[(I-N^{-1}(z^{-1},\theta))y_t+N^{-1}(z^{-1},\theta)M(z^{-1},\theta)r_t]\\&=N^{-1}(z^{-1},\theta)y_t-N^{-1}(z^{-1},\theta_0)M(z^{-1},\theta)r_t\\&=N^{-1}(z^{-1},\theta)[M(z^{-1},\theta_0)r_t+N(z^{-1},\theta_0)e_t]\\&\quad-N^{-1}(z^{-1},\theta)M(z^{-1},\theta)r_t\\&=N^{-1}(z^{-1},\theta)[M(z^{-1},\theta_0)-M(z^{-1},\theta)]r_t\\&\quad+N^{-1}(z^{-1},\theta)N(z^{-1},\theta_0)e_t\\&=N^{-1}(z^{-1},\theta)[M(z^{-1},\theta_0)-M(z^{-1},\theta)](e_t+\delta_t)\\&\quad+N^{-1}(z^{-1},\theta)N(z^{-1},\theta_0)e_t\end{aligned}$$

因此

$$\begin{aligned}\varepsilon(t,\theta)&=N^{-1}(z^{-1},\theta)[M(z^{-1},\theta_0)-M(z^{-1},\theta)]\delta_t\\&\quad+N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)+(M(z^{-1},\theta_0)-M(z^{-1},\theta))]e_t\end{aligned}$$

由于 δ_t 和 e_t 是独立的，所以有

$$\begin{aligned}\mathrm{Var}(\varepsilon(t,\theta))&=\mathrm{Var}(N^{-1}(z^{-1},\theta)[M(z^{-1},\theta_0)-M(z^{-1},\theta)]\delta_t)\\&\quad+\mathrm{Var}(N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)+(M(z^{-1},\theta_0)-M(Z^{-1},\theta))]e_t)\end{aligned} \tag{4.105}$$

此外，当 $a=a_0$、$b=b_0$ 时，使 $G_p(z^{-1},\theta)=G_p(z^{-1},\theta_0)$，从而使 $M(z^{-1},\theta)=M(z^{-1},\theta_0)$。其结果是，关于 δ_t 的方差项式（4.105）右侧是等于零的，因此

$$\mathrm{Var}(\varepsilon(t,\theta))=\mathrm{Var}(N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)+(M(z^{-1},\theta_0)-M(z^{-1},\theta))]e_t) \tag{4.106}$$

可以很容易地从给定的传递函数得到如下关系：

$$G_p(z^{-1}=0,\theta)=\frac{bz^{-2}}{1+az^{-1}}=0$$

$$G_p(z^{-1}=0,\theta_0)=\frac{b_0z^{-2}}{1+a_0z^{-1}}=0$$

$$G_l(z^{-1}=0,\theta)=\frac{1+cz^{-1}}{1+dz^{-1}}=1$$

$$G_l(z^{-1}=0,\theta_0)=\frac{1+c_0z^{-1}}{1+d_0z^{-1}}=1$$

$$C(z^{-1}=0)=\frac{k}{1-z^{-1}}=k$$

因此

$$M(z^{-1}=0,\theta_0)=\frac{G_p(z^{-1}=0,\theta_0)G_c(z^{-1}=0)}{1+G_p(z^{-1}=0,\theta_0)G_c(z^{-1}=0)}=0$$

$$M(z^{-1}=0,\theta)=\frac{G_p(z^{-1}=0,\theta)G_c(z^{-1}=0)}{1+G_p(z^{-1}=0,\theta)G_c(z^{-1}=0)}=0$$

$$N(z^{-1}=0,\theta_0)=\frac{G_l(z^{-1}=0,\theta_0)}{1+G_p(z^{-1}=0,\theta_0)G_c(z^{-1}=0)}=1$$

$$N(z^{-1}=0,\theta)=\frac{G_l(z^{-1}=0,\theta)}{1+G_p(z^{-1}=0,\theta)G_c(z^{-1}=0)}=1$$

当 $z^{-1}=0$ 时，其结果是

$$\lim_{z^{-1}\to 0}N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)+(M(z^{-1},\theta_0)-M(z^{-1},\theta))]e_t=e_t$$

因此，写作脉冲响应的形式，它可以很容易地证明：

$$N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)+(M(z^{-1},\theta_0)-M(z^{-1},\theta))]e_t$$
$$=e_t+\text{与 } e_{t-1},\ e_{t-2},\ \cdots\text{相关项}$$

由于 e_t 是独立于它过去的，从式（4.106）得出

$$\begin{aligned}\mathrm{Var}(\varepsilon(t,\theta))&=\mathrm{Var}(N^{-1}(z^{-1},\theta)[N(z^{-1},\theta_0)\\&\quad+(M(z^{-1},\theta_0)-M(z^{-1},\theta))]e_t)\\&\geqslant\mathrm{Var}(e_t)\end{aligned}$$

已经证明当 $a=a_0$、$b=b_0$ 时，$M(z^{-1},\ \theta)=M(z^{-1},\ \theta_0)$。当 $N(z^{-1},\ \theta)=N(z^{-1},\ \theta_0)$ 时，该值达到下界。因此，已经验证了在这些条件下用间接法估计过程和干扰模型的一致性。

4.6　子空间辨识

由于 PEM 是一种基于优化数值的辨识方法，它自然存在可能的局部最优问题。其计算尤其对多输入多输出系统相对复杂。另一方面，子空间识别法提供基于计算工具的替代解决方案，例如作为 QR 因子分解和 SVD（奇异值分解），这使得它们从一个数值点本质增强。子空间辨识法也是非迭代并避免局部极小，且没有任何收

敛问题。如它是基于状态空间模型的结构，子空间辨识方法本质上适合于多变量系统辨识。不过，这些方法的表现通常不如 PEM，因为它们不能达到如 PEM 求出的最大似然估计一样。

4.6.1　符号

考虑对于线性系统的状态空间表示法的创新形式，其输入 l 且输出 m 如下：

$$\begin{cases} x_{t+1} = Ax_t + Bu_t + Ke_t \\ y_t = Cx_t + Du_t + e_t \end{cases} \tag{4.107}$$

式中，$x_t \in \mathbb{R}^n$；$u_t \in \mathbb{R}^l$；$y_t \in \mathbb{R}^m$；$e_t \in \mathbb{R}^m$。

子空间识别的第一步是形成为一个有限时间范围状态空间方程表示的矩阵的状态/输出轨迹。出于这个原因，需要在有限的时间范围解式（4.107）。

根据式（4.107），当 $t=N$ 时有

$$x_{N+1} = Ax_N + Bu_N + Ke_N \tag{4.108}$$

当 $t=N+1$ 时有

$$\begin{aligned} x_{N+2} &= Ax_{N+1} + Bu_{N+1} + Ke_{N+1} \\ &= A^2 x_N + (AB \quad A)\begin{pmatrix} u_N \\ u_{N+1} \end{pmatrix} + (AK \quad K)\begin{pmatrix} e_N \\ e_{N+1} \end{pmatrix} \end{aligned}$$

这个步骤继续到 $t=2N-2$：

$$\begin{aligned} x_{2N-1} = A^{N-1}x_N &+ (A^{N-2}B \quad A^{N-3}B \quad \cdots B)\begin{pmatrix} u_N \\ u_{N+1} \\ \vdots \\ u_{2N-2} \end{pmatrix} \\ &+ (A^{N-2}K \quad A^{N-3}K \quad \cdots K)\begin{pmatrix} e_N \\ e_{N+1} \\ \vdots \\ e_{2N-2} \end{pmatrix} \end{aligned} \tag{4.109}$$

考虑输出式（4.107），可以得出以下关系：

$$\begin{aligned} y_{2N-1} = CA^{N-1}x_N &+ (CA^{N-2}B \quad CA^{N-3}B \ \cdots D)\begin{pmatrix} u_N \\ u_{N+1} \\ \vdots \\ u_{2N-1} \end{pmatrix} \\ &+ (CA^{N-2}K \quad CA^{N-3}K \quad \cdots I)\begin{pmatrix} e_N \\ e_{N+1} \\ \vdots \\ e_{2N-1} \end{pmatrix} \end{aligned} \tag{4.110}$$

把 $t=N$，$N+1$，…，$2N-1$ 结果放在一起得出状态空间方程的矩阵解，如下：

$$
\begin{pmatrix} y_N \\ y_{N+1} \\ y_{N+2} \\ \vdots \\ y_{2N-1} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{pmatrix} x_N
+ \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix} \begin{pmatrix} u_N \\ u_{N+1} \\ u_{N+2} \\ \vdots \\ u_{2N-1} \end{pmatrix}
+ \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix} \begin{pmatrix} e_N \\ e_{N+1} \\ e_{N+2} \\ \vdots \\ e_{2N-1} \end{pmatrix} \tag{4.111}
$$

增加时间指数到 1，式（4.111）变为

$$
\begin{pmatrix} y_{N+1} \\ y_{N+2} \\ y_{N+3} \\ \vdots \\ y_{2N} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{pmatrix} x_{N+1}
+ \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix} \begin{pmatrix} u_{N+1} \\ u_{N+2} \\ u_{N+3} \\ \vdots \\ u_{2N} \end{pmatrix}
+ \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix} \begin{pmatrix} e_{N+1} \\ e_{N+2} \\ e_{N+3} \\ \vdots \\ e_{2N} \end{pmatrix} \tag{4.112}
$$

继续这个过程，添加时间指数 2 和 3 直到 $j-1$ 并把结果方程放到一起以矩阵形式给出：

$$
\begin{pmatrix} y_N & y_{N+1} & \cdots & y_{N+j-1} \\ y_{N+1} & y_{N+2} & \cdots & y_{N+j} \\ y_{N+2} & y_{N+3} & \cdots & y_{N+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{2N-1} & y_{2N} & \cdots & y_{2N+j-2} \end{pmatrix}
= \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{pmatrix} (x_N \quad x_{N+1} \cdots x_{N+j-1})
+ \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & D \end{pmatrix}
\begin{pmatrix} u_N & u_{N+1} & \cdots & u_{N+j-1} \\ u_{N+1} & u_{N+2} & \cdots & u_{N+j} \\ u_{N+2} & u_{N+3} & \cdots & u_{N+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ u_{2N-1} & u_{2N} & \cdots & u_{2N+j-2} \end{pmatrix}
+ \begin{pmatrix} I & 0 & \cdots & 0 \\ CK & I & \cdots & 0 \\ CAK & CK & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}K & CA^{N-3}K & \cdots & I \end{pmatrix}
\begin{pmatrix} e_N & e_{N+1} & \cdots & e_{N+j-1} \\ e_{N+1} & e_{N+2} & \cdots & e_{N+j} \\ e_{N+2} & e_{N+3} & \cdots & e_{N+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ e_{2N-1} & e_{2N} & \cdots & e_{2N+j-2} \end{pmatrix} \tag{4.113}
$$

式（4.113）为系统提供输出的子空间方程。在状态方程执行类似的程序，对系统状态给出了下面的子空间方程：

$$
(x_N \quad x_{N+1} \cdots \quad x_{N+j-1}) = A^N (x_0\ x_1 \cdots x_{j-1})
+ (A^{N-1}B \quad A^{N-2}B \cdots B) \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \vdots & \vdots \\ u_{N-1} & u_N & \cdots & u_{N+j-2} \end{pmatrix}
+ (A^{N-1}K \quad A^{N-2}K \cdots K) \begin{pmatrix} e_0 & e_1 & \cdots & e_{j-1} \\ e_1 & e_2 & \cdots & e_j \\ \vdots & \vdots & \vdots & \vdots \\ e_{N-1} & e_N & \cdots & e_{N+j-2} \end{pmatrix} \tag{4.114}
$$

表示式（4.113）用简单的记法：

$$
\boldsymbol{Y}_f = \boldsymbol{\Gamma}_N \boldsymbol{X}_f + \boldsymbol{L}_u \boldsymbol{U}_f + \boldsymbol{L}_e \boldsymbol{E}_f
$$

记式（4.114）为

$$\boldsymbol{X}_f = \boldsymbol{A}^N \boldsymbol{X}_p + \boldsymbol{\Delta}^d \boldsymbol{U}_p + \boldsymbol{\Delta}^s \boldsymbol{E}_p$$

现在，如果从 $t=0$ 开始并以相同的方法推导式（4.113），可得出下面的子空间方程：

$$\begin{pmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_j \\ y_2 & y_3 & \cdots & y_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{N-1} & y_N & \cdots & y_{N+j-2} \end{pmatrix}$$

$$= \begin{pmatrix} C \\ CA \\ CA^2 \\ CA^{N-1} \end{pmatrix} (x_0\ x_1 \cdots x_{j-1})$$

$$+ \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & D \end{pmatrix} \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ u_2 & u_3 & \cdots & u_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ u_{N-1} & u_N & \cdots & u_{N+j-2} \end{pmatrix}$$

$$+ \begin{pmatrix} I & 0 & \cdots & 0 \\ CK & I & \cdots & 0 \\ CAK & CK & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}K & CA^{N-3}K & \cdots & I \end{pmatrix} \begin{pmatrix} e_0 & e_1 & \cdots & e_{j-1} \\ e_1 & e_2 & \cdots & e_j \\ e_2 & e_3 & \cdots & e_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ e_{N-1} & e_N & \cdots & e_{N+j-2} \end{pmatrix} \qquad (4.115)$$

式（4.115）表示为

$$\boldsymbol{Y}_p = \boldsymbol{\Gamma}_N \boldsymbol{X}_p + \boldsymbol{L}_n \boldsymbol{U}_p + \boldsymbol{L}_e \boldsymbol{E}_p$$

把下面 3 个简写子空间方程放在一起给出：

$$\boldsymbol{Y}_f = \boldsymbol{\Gamma}_N \boldsymbol{X}_f + \boldsymbol{L}_u \boldsymbol{U}_f + \boldsymbol{L}_e \boldsymbol{E}_f \qquad (4.116)$$

$$\boldsymbol{Y}_p = \boldsymbol{\Gamma}_N \boldsymbol{X}_p + \boldsymbol{L}_u \boldsymbol{U}_p + \boldsymbol{L}_e \boldsymbol{E}_p \qquad (4.117)$$

$$\boldsymbol{X}_f = \boldsymbol{A}^N \boldsymbol{X}_p + \boldsymbol{\Delta}^d \boldsymbol{U}_p + \boldsymbol{\Delta}^s \boldsymbol{E}_p \qquad (4.118)$$

式中，下标“p”代表“过去”；“f”代表“未来”。

比较式（4.116）与式（4.113）定义子空间矩阵 $\boldsymbol{\Gamma}_N$、$\boldsymbol{L}_u$ 和 $\boldsymbol{L}_e$ 为

$$\boldsymbol{\Gamma}_N = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{pmatrix} \qquad (4.119)$$

$$\boldsymbol{L}_u = \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix} \tag{4.120}$$

$$\boldsymbol{L}_e = \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix} \tag{4.121}$$

式中，$\boldsymbol{\Gamma}_N \in \mathbb{R}^{mN \times n}$为扩展可观性矩阵；$\boldsymbol{L}_u \in \mathbb{R}^{mN \times lN}$和$\boldsymbol{L}_e \in \mathbb{R}^{mN \times mN}$分别为包含过程和扰动模型马尔科夫（Markov）参数。

数据 Hanlke 矩阵 $\boldsymbol{U}_p$、$\boldsymbol{U}_f \in \mathbb{R}^{lN \times j}$、$\boldsymbol{Y}_p$、$\boldsymbol{Y}_f \in \mathbb{R}^{mN \times j}$和$\boldsymbol{E}_p$、$\boldsymbol{E}_f \in \mathbb{R}^{mN \times j}$可以通过比较式（4.116）和式（4.113）、式（4.117）和式（4.115）来定义。例如，$\boldsymbol{U}_p$和 $\boldsymbol{U}_f$ 定义如下：

$$\boldsymbol{U}_p = \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \vdots & \vdots \\ u_{N-1} & u_N & \cdots & u_{N+j-2} \end{pmatrix}$$

$$\boldsymbol{U}_f = \begin{pmatrix} u_{\boldsymbol{N}} & u_{\boldsymbol{N}+1} & \cdots & u_{\boldsymbol{N}+\boldsymbol{j}-1} \\ u_{\boldsymbol{N}+1} & u_{\boldsymbol{N}+2} & \cdots & u_{\boldsymbol{N}+\boldsymbol{j}} \\ \vdots & \vdots & \vdots & \vdots \\ u_{2\boldsymbol{N}-1} & u_{2\boldsymbol{N}} & \cdots & u_{2\boldsymbol{N}+\boldsymbol{j}-2} \end{pmatrix}$$

其他矩阵也可以类似地建立。一般地，j 要比 max（mN，lN）大，以减少对噪声的敏感度（Overschee 和 Moor，1996）。在所有的子空间辨识方法中，有一些常见的假设，如数对（A，C）是可观察到的与数对（A，$BK\boldsymbol{R}^{1/2}$）应该是可控的，其中 $\boldsymbol{R}$ 是白噪声在输入过程的协方差矩阵。

4.6.2 回归分析法子空间辨识

下列推导来自 Knudsen（2001）的步骤。再次考虑在式（4.107）中的空间模型状态。通过转化时间指数得到：

$$x_t = Ax_{t-1} + Bu_{t-1} + Ke_{t-1}$$
$$y_{t-1} = Cx_{t-1} + Du_{t-1} + e_{t-1}$$

合并以上方程得到

$$x_t = (A - KC)x_{t-1} + Ky_{t-1} + (B - KD)u_{t-1}$$

递归相同替换得到

$$x_t = (A - KC)^N x_{t-N} + (A - KC)^{N-1}(Ky_{t-N} + (B - KD)u_{t-N}) + \cdots + (A - KC)(Ky_{t-2} + (B - KD)u_{t-2}) + Ky_{t-1} + (B - KD)u_{t-1} \tag{4.122}$$

式（4.122）的时间指数可以通过 $t = N$，$N+1$，…，$N+j-1$ 替换。综合所有结论方程以矩阵形式给出一个子空间方程，如下：

$$\boldsymbol{X}_f = \boldsymbol{\Phi}_y \boldsymbol{Y}_p + \boldsymbol{\Phi}_u \boldsymbol{U}_p + \boldsymbol{\Phi}_x \boldsymbol{X}_p \tag{4.123}$$

式中

$$\boldsymbol{\Phi}_x = (A - KC)^N$$

$$\boldsymbol{\Phi}_y = C(A - KC, K)$$

$$\boldsymbol{\Phi}_u = C(A - KC, B - KD)$$

并且将 C 操作定义为

$$C(A,B) \triangleq (A^{N-1}B \quad A^{N-2}B \quad \cdots \quad AB \quad B) \tag{4.124}$$

Φ_x 代表卡尔曼滤子的误差动态。当 N→∞ 时，有 $\boldsymbol{\Phi}_x \to 0$，因为卡尔曼滤子的稳定性，结果得到大 N：

$$\boldsymbol{X}_f = \boldsymbol{\Phi}_y \boldsymbol{Y}_p + \boldsymbol{\Phi}_u \boldsymbol{U}_p \tag{4.125}$$

用式（4.116）代替式（4.125）得到

$$\boldsymbol{Y}_f = \boldsymbol{L}_w \boldsymbol{W}_p + \boldsymbol{L}_u \boldsymbol{U}_f + \boldsymbol{L}_e \boldsymbol{E}_f \tag{4.126}$$

其中超越数据矩阵 $\boldsymbol{W}_p$ 定义为

$$\boldsymbol{W}_p = \begin{pmatrix} \boldsymbol{Y}_p \\ \boldsymbol{U}_p \end{pmatrix} \tag{4.127}$$

$\boldsymbol{L}_w$ 可以表示为

$$\boldsymbol{L}_w = \boldsymbol{\Gamma}_N (\boldsymbol{\Phi}_y \quad \boldsymbol{\Phi}_y) \tag{4.128}$$

在式（4.126）的基础上，子空间矩阵可以通过最小二乘法估计：

$$\begin{aligned} (\hat{\boldsymbol{L}}_w \quad \hat{\boldsymbol{L}}_u) &= \boldsymbol{Y}_f \begin{pmatrix} \boldsymbol{W}_p \\ \boldsymbol{U}_f \end{pmatrix}^{\dagger} \\ &= \boldsymbol{Y}_f (\boldsymbol{W}_P^T \quad \boldsymbol{U}_f^T) \left(\begin{pmatrix} \boldsymbol{W}_p \\ \boldsymbol{U}_f \end{pmatrix} (\boldsymbol{W}_P^T \quad \boldsymbol{U}_f^T) \right)^{-1} \end{aligned} \tag{4.129}$$

这个伪逆操作可以用 QR 因式分解跟数字化鲁棒法完成。这个最小二乘法估计的残差给出如下：

$$\boldsymbol{V}_f = \boldsymbol{Y}_f - \hat{\boldsymbol{Y}}_f \tag{4.130}$$

$$= \boldsymbol{L}_e \hat{\boldsymbol{E}}_f \tag{4.131}$$

它的协方差可以估计为

$$\hat{\boldsymbol{P}}_v = \frac{1}{j} \boldsymbol{V}_f \boldsymbol{V}_f^{\mathrm{T}} \tag{4.132}$$

进一步，创新项 $\boldsymbol{e}_f$ 可从 $\boldsymbol{V}_f$ 第一行估计为如下：

$$\hat{\boldsymbol{e}}_f = \boldsymbol{V}_f(1:m,:) \tag{4.133}$$

为了检索系统矩阵，需要进行如下 *SVD*：

$$\boldsymbol{W}_1\hat{\boldsymbol{L}}_w\boldsymbol{W}_2 = (\boldsymbol{U}_1 \quad \boldsymbol{U}_2)\begin{pmatrix}\boldsymbol{S}_1 & 0\\ 0 & \boldsymbol{S}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{V}_1\\ \boldsymbol{V}_2\end{pmatrix} \tag{4.134}$$

式中，$\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 是真加权矩阵。

Knudsen（2001）给出了加权矩阵：

$$\boldsymbol{W}_1 = \boldsymbol{I} \tag{4.135}$$

$$\boldsymbol{W}_2 = (\boldsymbol{W}_p\boldsymbol{W}_p^{\mathrm{T}})^{1/2} \tag{4.136}$$

理论上，对于一个有限维数的状态空间模型，存在 $\boldsymbol{S}_2=0$，并且 $\boldsymbol{S}_1$ 的维数决定了状态空间矩阵 $\boldsymbol{A}$ 的维数。从这个分解，$\boldsymbol{\Gamma}_N$ 可以通过以下方式得到：

$$\hat{\boldsymbol{\Gamma}}_N = \boldsymbol{U}_1\boldsymbol{S}_1^{1/2} \tag{4.137}$$

系统矩阵现在可以检索如下（基于 MATLAB ® 注释）：

$$\hat{\boldsymbol{\Gamma}}_N^1 = \hat{\boldsymbol{\Gamma}}_N(1:(N-1)m,:) \tag{4.138}$$

$$\hat{\boldsymbol{\Gamma}}_N^2 = \hat{\boldsymbol{\Gamma}}_N(m+1:Nm,:) \tag{4.139}$$

$$\hat{\boldsymbol{C}} = \hat{\boldsymbol{\Gamma}}_N(1:m,1:n) \tag{4.140}$$

$$\hat{\boldsymbol{A}} = (\hat{\boldsymbol{\Gamma}}_N^1)^{\dagger}\hat{\boldsymbol{\Gamma}}_N^2 \tag{4.141}$$

$$\hat{\boldsymbol{B}} = (\boldsymbol{\Gamma}_N^1)^{\dagger}\hat{\boldsymbol{L}}_u(m+1:Nm,1:l) \tag{4.142}$$

$$\hat{\boldsymbol{D}} = \hat{\boldsymbol{L}}_u(1:m,1:l) \tag{4.143}$$

$$\hat{\boldsymbol{K}} = (\hat{\boldsymbol{\Gamma}}_N^1)_{\dagger}\hat{\boldsymbol{P}}_v(m+1:Nm,1:m)\hat{\boldsymbol{R}}^{-1} \tag{4.144}$$

在子空间辨识中，还有一些基于其他的方法，例如基于投影的方法（例如数值算法用于子空间状态空间系统辨识或 N4SID）或统计方法（如典型变量分析或 CVA）和多变量输出误差状态空间（MOESP）方法。这些方法的概述可以在 Huang 和 Kadali（2008）的文章和其中的参考文献中找到。

许多方法也已经被开发为基于闭环系统辨识子空间的方法。改进 N4SID、ARX 的预测方法、创新估算法、正交投影方法和联合输入输出辨识是一些例子。闭环子空间辨识方法的概述可以在 Huang 和 Kadali（2008）的文章和其中的参考文献中找到。

为了说明这一点，实施回归分析的一个简单的算例方法子空间辨识在下面的内容中给出。

4.6.3 范例

考虑下面的状态空间模型：

$$x(t+1) = -0.1x(t) + u(t) + 0.2e(t) \tag{4.145}$$

$$y(t) = 0.5x(t) + e(t) \tag{4.146}$$

噪声方差设定为0.01。为了便于说明，一个小的组输入输出数据见表4.1，是在的这个过程中产生的。

表4.1　验证数据

$u(t)$	1	−1	1	−1	−1	1
$y(t)$	0.0021	0.5002	−0.5479	0.5562	−0.5536	−0.4435

选择 $N=2$ 且 $j=3$，数据 Hankel 矩阵可以构造如下：

$$\boldsymbol{U}_p=\begin{pmatrix}1 & -1 & 1\\ -1 & 1 & -1\end{pmatrix} \tag{4.147}$$

$$\boldsymbol{U}_f=\begin{pmatrix}1 & -1 & -1\\ -1 & -1 & 1\end{pmatrix} \tag{4.148}$$

$$\boldsymbol{Y}_p=\begin{pmatrix}0.0021 & 0.5002 & -0.5479\\ 0.5002 & -0.5479 & 0.5562\end{pmatrix} \tag{4.149}$$

$$\boldsymbol{Y}_f=\begin{pmatrix}-0.5479 & 0.5562 & -0.5536\\ 0.5562 & -0.5536 & -0.4435\end{pmatrix} \tag{4.150}$$

然后，$\boldsymbol{W}_p$ 由下式给出：

$$\boldsymbol{W}_p=\begin{pmatrix}0.0021 & 0.5002 & -0.5479\\ 0.5002 & -0.5479 & 0.5562\\ 1 & -1 & 1\\ -1 & 1 & -1\end{pmatrix} \tag{4.151}$$

用式（4.129）可以估计 $\boldsymbol{L}_w$ 和 $\boldsymbol{L}_u$ 如下：

$$\begin{aligned}(\hat{\boldsymbol{L}}_w\hat{\boldsymbol{L}}_u) &= \begin{pmatrix}-0.5479 & 0.5562 & -0.5536\\ 0.5562 & -0.5536 & -0.4435\end{pmatrix}\\ &\times\begin{pmatrix}0.0021 & 0.5002 & -0.5479\\ 0.5002 & -0.5479 & 0.5562\\ 1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & -1\\ -1 & -1 & 1\end{pmatrix}^{\dagger}=\begin{pmatrix}-0.5479 & 0.5562 & -0.5536\\ 0.5562 & -0.5536 & -0.4435\end{pmatrix}\\ &\times\begin{pmatrix}0.0021 & 0.5002 & 1 & -1 & 1 & -1\\ 0.5002 & -0.5479 & -1 & 1 & -1 & -1\\ -0.5479 & 0.5562 & 1 & -1 & -1 & 1\end{pmatrix}\end{aligned}$$

$$\times\left(\begin{pmatrix}0.0021 & 0.5002 & -0.5479\\ 0.5002 & -0.5479 & 0.5562\\ 1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & -1\\ -1 & -1 & 1\end{pmatrix}\right.$$

$$\left.\times\begin{pmatrix}0.0021 & 0.5002 & 1 & -1 & 1 & -1\\ 0.5002 & -0.5479 & -1 & 1 & -1 & -1\\ -0.5479 & 0.5562 & 1 & -1 & -1 & 1\end{pmatrix}\cdot\right)^{-1}$$

$$=\begin{pmatrix}0.0604 & -0.1235 & -0.2345 & 0.2345 & -0.0032 & 0.0140\\ 0.0049 & 0.0116 & 0.0258 & -0.0258 & 0.4985 & -0.0003\end{pmatrix}\tag{4.152}$$

得出：

$$\hat{\boldsymbol{L}}_w=\begin{pmatrix}0.0604 & -0.1235 & -0.2345 & 0.2345\\ 0.0049 & 0.0116 & 0.0258 & -0.0258\end{pmatrix}$$

$$\hat{\boldsymbol{L}}_u=\begin{pmatrix}-0.0032 & 0.0140\\ 0.4985 & -0.0003\end{pmatrix}\tag{4.153}$$

式（4.130）能用来估计 $\boldsymbol{V}_f$ 为

$$\boldsymbol{V}_f=\begin{pmatrix}0.0142 & -0.0172 & 0.0126\\ -0.1524 & 0.0186 & 0.2112\end{pmatrix}\tag{4.154}$$

得出：

$$\hat{\boldsymbol{P}}_v=\begin{pmatrix}0.0002 & 0.00006\\ 0.00006 & 0.0226\end{pmatrix}$$

$$\hat{\boldsymbol{e}}_f=\boldsymbol{V}_f(1,:)=(0.0142\quad -0.0172\quad 0.0126)$$

$$\hat{\boldsymbol{R}}=\mathrm{cov}(\boldsymbol{e}_f)=0.0003\tag{4.155}$$

SVD 在 $\hat{\boldsymbol{L}}_w$ 下得出：

$$\hat{\boldsymbol{L}}_w=\begin{pmatrix}0.9948 & -0.1022\\ -0.1022 & 0.9948\end{pmatrix}\begin{pmatrix}0.3609 & 0\\ 0 & 0.0114\end{pmatrix}\begin{pmatrix}-0.1651 & 0.9727\\ 0.3437 & -0.0982\end{pmatrix}$$

使用式（4.137），$\boldsymbol{\Gamma}_N$ 可以估计如下：

$$\hat{\boldsymbol{\Gamma}}_N=\begin{pmatrix}0.9948\\ -0.1022\end{pmatrix}(0.3609)^{1/2}=\begin{pmatrix}0.5976\\ -0.0614\end{pmatrix}\tag{4.156}$$

现在，该系统参数 $\{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},\boldsymbol{k}\}$ 可利用式（4.140）~式（4.144）来估计：

$$\hat{\boldsymbol{c}}=\hat{\boldsymbol{\Gamma}}_N(1,1)=0.5976\tag{4.157}$$

$$\hat{\boldsymbol{a}}=(\boldsymbol{\Gamma}_N^u)^{\dagger}\boldsymbol{\Gamma}_N^y=(\hat{\boldsymbol{c}})^{-1}\hat{\boldsymbol{\Gamma}}_N(2,1)=(0.5976)^{-1}(-0.0614)=-0.1228\tag{4.158}$$

$$\hat{\boldsymbol{b}}=(\boldsymbol{\Gamma}_N^u)^{\dagger}\hat{\boldsymbol{L}}_u(2,1)=(\hat{\boldsymbol{c}})^{-1}\boldsymbol{L}_{\hat{u}}(2,1)=(0.5976)^{-1}(0.4995)=0.9969\tag{4.159}$$

$$\hat{d} = \hat{L}_u(1,1) = -0.0032 \tag{4.160}$$

$$\hat{k} = (\Gamma_N^u)^{\dagger}\hat{P}(2,1)(\hat{R})^{-1} = (0.5976)^{-1}(0.00006)(0.0003)^{-1} = 0.3347 \tag{4.161}$$

$$\hat{R} = 0.0003 \tag{4.162}$$

式（4.161）和式（4.162）显示噪声动态和噪声方差较差的估计，它可以作为使用小数目输入输出数据的结果。

4.7　注解和参考文献

系统辨识的经典理论已由 Söderström 和 Stoica（1989）以及 Ljung（1999）进行了详细讨论。近来，在应用方面 Huang 和 Kadali（2008）做出了一定贡献（Springer-Verlag）。本章为系统辨识理论和应用的教程。本章的主要参考资料为 Söderström 和 Stoica（1989）、Ljung（1999）、Huang 和 Kadali（2008）的研究（Springer-Verlag）。PEM 是在实践中应用最广泛的辨识方法，其他受欢迎的辨识方法包括工具变量法，工具变量法的详细讨论可以在 Söderström 和 Stoica（1989）的研究中找到。对于本章中介绍的子空间法主要参考是 Daneshpour（2009）和 Huang、Kadali（2008）的研究（Springer-Verlag）。深入讨论子空间法可在 Overschee、Moor（1996）和 Katayama（2005）的研究中找到。

本章介绍的例题是阿尔伯塔大学系统辨识课程提供的第一作者讲义，一些例题的解是出于 Khatibisepehr 等人的解答（2009-2011）。

第 5 章　状 态 估 计

5.1　随机动态系统过滤技术进展

一旦一个过程模型是有效的，无论它是以基本原理为基础的模型还是以数据为基础的模型，它都被应用于多方面，例如仿真模拟、状态估计和控制。本章重点关注的是状态估计问题。

状态估计与作为最具代表性的卡尔曼滤波（KF）（kalman，1960）对于线性随机动态系统已经成为一个相对成熟的项目。然而，状态估计对于非线性随机动态系统在最近几年以前一直惊人地被限制。多年以来，扩展卡尔曼滤波（EKF）（Jazwinski，1970）似乎是唯一的选择。在 20 世纪 90 年代，多种多样的滤波和状态估计技术发展起来。最近几年中，新兴的贝叶斯方法因其计算能力容量的增加受到越来越多的注意。在最近的发展中，最值得注意的是无味卡尔曼滤波（UKF）（Julier 和 Uhlmann，1997，2004；Norgaard 等人，2000；Wan 和 van der Merwe，2001；Wan 等人，2000），集合卡尔曼滤波（EnKF）（Evensen，1994，2009）和粒子滤波（Doucet 等人，2001；Gordon 等人，1993）。UKF 的成功应用被报道，甚至用于混乱系统（Sitz 等人，2002）。

贝叶斯状态估计方法的目标是用所有可用的测量估计一种状态后面的概率密度函数（pdf）。用字母 p（状态/证据）表示，运算符 p 或者概率代表了根据现有证据确定状态的不确定性。由于更多的证据变得容易获得，不确定性普遍降低。基于这些，贝叶斯方法通常被用于在线状态估计来利用当前所有证据和递归当前继续进行的状态估计过程作为新证据成为可能。一旦 p（状态/证据）被计算，根据确定的优化标准，例如最小方差误差或者最大后验概率（MAP）原理，一个状态的具体点估计能够达到。

贝叶斯解决方案提供了一个完整的概率密度函数，而且能够处理多模态、非对称和不连续（Julier 和 Uhlmann，2004）。这与通常采取高斯概率和线性参数的经典估值形成对比。然而并不是所有的概率密度函数可以用有限数量的参数来描述（高斯概率仅需要两个参数，即平均值和方差）。因而，对于一般的贝叶斯推理问题的解决方案必须依赖概率密度函数的近似值，因为高计算负荷是普遍预期的。

一些概率密度函数的近似值在相关文献中被提出。不幸的是，它们中的大多数要么计算难以管理，要么需要在实践中不容易满足过程的形式和观察模型的特殊假设。需要均值和方差作为充足矩量的 KF 来代替概率密度函数依然是最受欢迎的方

法。EKF（Jazwinski，1970）对于非线性系统大概是应用最广泛的估算方法。在Julier和合作者的一系列论文中已经表明常规EKF在估算均值和协方差时仅仅精确到一阶，并且需要Jacobian行列式的计算。作为一个进步，二阶EKF已经被提出，但是由于对Hessian矩阵的需要，它的应用仍需要大量的时间努力。

过滤的另外一种直观的方法是直接由一组典型采样来估计概率密度函数。对这些采样应用状态空间方程以产生新的预测样本，从中所需要的统计，如平均值和协方差，可估计并用于KF更新方程（Sitz等人，2002）。沿着这个思路，Julier和Uhlmann对非线性系统提出了一种新的扩展的KF，称为EKF。这个程序属于其中的概率密度函数被截断（Sitz等人，2002）的统计线性化方案类别。密度函数的高阶矩被忽视，也就是说只有平均值和协方差被使用。一组具有相同平均值和协方差的采样点产生并且通过充分的状态空间模型传播。与EKF相比，UKF处理非线性问题用一种更为简单但是优雅的方式，严格意义上来说，即一种相似的或者与EKF相比不增加计算量来达到更好的估计。它还表明了UKF对平均值和协方差的估计精确到二阶（Julier和Uhlmann，2004）。不依赖于Jacobian行列式的导数，UKF能够处理不连续问题。由于UKF的特殊结构，停止在一个局部最小值的成本函数的概率减小，并且它也可以被应用到不均匀采样数据。

然而，一个趋近概率密度函数的更全面的方法是通过颗粒过滤（Doucet等人，2001；Gordon等人，1993）。颗粒是从概率密度函数中提取的随机采样。一个颗粒过滤本质上是一个蒙特卡罗模拟（Monte Carlo Simulator），目的是通过递归的趋近完整的概率分布函数直接完成贝叶斯估计（van der Merwe，2004）。后验概率密度函数是由一组对其形式不做明确假设的加权采样粗略估计的，因此可以在一般的非线性/非高斯系统中使用。详细点评和颗粒过滤的工程应用的讨论可在最近的许多参考文献中找到，例如Chen等人（2004）和Shao等人（2011）的研究。

状态估计问题的解决方案可以超越传统的“状态”。众所周知的是，参数估计问题可以表述为状态估计问题（Nelson，2000）。这种相同的方法继续存在在UKF中（Gustafsson和Hriljac，2003；van der Merwe，2004；Wan和van der Merwe，2000）。本章考虑状态估计以及在一个统一的框架下随机非线性动态系统的参数估计。在第3章和第4章中，模型参数是通过黑箱方法估计的。在本章中，参数被估计为具有已知模型结构模型。因此，本章中介绍的参数估计方法将用来估计第一性原理模型的未知参数或者给定模型的未知参数。为了这个目的，本章首先讨论了在状态估计的框架下，参数估计问题是如何表达的。然后，介绍了序列贝叶斯推理方法，这提供了一个对状态估计问题通用的解决方案。检查了KF和EKF的贝叶斯解释以后，专注于UKF对状态和参数的估计。在本章中，竭力用简单的语言去解释数学概念，并提供了一些例子来阐明UKF对非线性系统状态和参数估计的应用。

5.2 问题界定

动态系统的参数估计的最根本的方法是黑盒方法。除了黑盒方法，也有白盒方法和灰盒方法。按照 Lindskog（1996）的说法，白盒模型被定义为反映真实系统所有特性的一种模型。它们仅仅是由原有的知识和物理原理构建，而没有系统中任何有效的测量。这种类型的模型技术也被称为在前面的章节中介绍的第一性原理模型。

然而，如果一个人追求白盒模型，但是调整某些或者全部物理参量，那么他实际上是在做参数估计。这种类型的建模也通常称为灰盒模型，也就是说，底层系统的确定模型结构是根据第一性原理得出的，但是参数却是未知的。灰盒非线性动态系统模型就系统的动态未知而受到挑战。从时间序列估计参数化模型的问题的一般方法产生状态空间描述（Sitz 等人，2002）。但即使是线性状态空间模型，模型参数的估计成为一个非线性估计问题，因为这两个状态和参数都是未知的。

下面通过开始一个静态非线性模型开始对灰盒模型的参数估计进行讨论：

$$y_t = g\ (u_t,\ \theta,\ e_t)$$

式中，u_t 是输入；y_t 是输出，θ 是一组未知参数。

给出一组 u_t 和 y_t 的数据，参数估计的目的是估计出最小化残留的 θ 值：

$$\varepsilon_t = y_t - \hat{g}(u_t, \theta) \tag{5.1}$$

在一个特定的最佳方式。

这个问题可以用一个状态估计问题表达为

$$\theta_{t+1} = \theta_t$$
$$y_t = g\ (u_t,\ \theta,\ e_t)$$

这一公式假设参数 θ 的值是恒定的，如果 θ 也是随时间变化的，那么问题可以用公式表达成：

$$\theta_{t+1} = \theta_t + w_t$$
$$y_t = g\ (u_t,\ \theta,\ e_t)$$

式中，w_t 的分布决定了参数随着时间变化。

但是，即使参数是恒定的，但仍然需要添加的噪声项。要启动参数估计，初始猜测值总是与实际不相同的。在无噪声项的情况下，参数估计会得到与初始猜测值同样的数值。因而，w_t 的分布可以被用来作为参数分布的调谐旋钮。例如，w_t 的大方差的选择导致了参数的快速跟踪但是导致参数估计更多的差异。w_t 的微小差异会是估计顺利，但是减慢可变参数的跟踪。这一提法的实质是要考虑 θ_t 作为状态，这样的话参数估计问题成为一个状态估计问题。现在，考虑一个非线性状态空间表示的非线性动态系统：

$$z_{t+1} = f(z_t, u_t, \theta_t) + v_t^z \tag{5.2}$$

$$y_t = h(z_t, u_t, \theta_t) + e_t \tag{5.3}$$

式中，$z_t \in \mathbb{R}^n$是状态；$\theta_t \in \mathbb{R}^d$是参量；$v_t^z \in \mathbb{R}^n$是干扰；$u_t \in \mathbb{R}^q$是输入；$y_t \in \mathbb{R}^m$是输出。

按照和静态模型参数估计同样的方法，状态和非线性动态模型的参数估计问题可归结为

$$\begin{bmatrix} z_{t+1} \\ \theta_{t+1} \end{bmatrix} = \begin{bmatrix} f(z_t, u_t, \theta_t) \\ \theta_t \end{bmatrix} + \begin{bmatrix} v_t^z \\ v_t^\theta \end{bmatrix} \tag{5.4}$$

$$y_t = h(z_t, u_t, \theta_t) + e_t \tag{5.5}$$

令

$$x_t = \begin{bmatrix} z_t \\ \theta_t \end{bmatrix}$$

然后参数和状态的估计可归结为 x_t 的估计与附加噪声下非线性动态模型：

$$x_{t+1} = f\ (x_t,\ u_t)\ + v_t$$

$$y_t = h\ (x_t,\ u_t)\ + e_t$$

Gustafsson 和 Hriljac（2003）提出了另一个可供选择的构想，其中增广状态空间模型可以写成：

$$\begin{bmatrix} z_{t+1} \\ \theta_{t+1} \end{bmatrix} = \begin{bmatrix} f(z_t, u_t, \theta_t) \\ \theta_t \end{bmatrix} + \begin{bmatrix} v_t^z + w_t^z \\ v_t^\theta + w_t^\theta \end{bmatrix} \tag{5.6}$$

$$y_t = h(z_t, u_t, \theta_t) + e_t \tag{5.7}$$

式中，噪声项包括物理状态的噪声；v_t^z 代表参数的不确定性的噪声项；v_t^θ、w_t^z 是状态粗化噪声；w_t^θ 是参数粗化噪声，粗化噪声是第二级的调谐参数。

例如，在时不变系统的系统辨识下，其中 $v_t^\theta = 0$，w_t^θ 应选择具有方差衰减至零，以保证参数估计的收敛性。

5.3　状态估计的序贯贝叶斯推理

现在，参数估计问题已经被表达成状态估计问题，聚焦于结果上的状态估计。设想一般状态模型代表非线性动态系统：

$$x_{t+1} = f(x_t, v_t) \tag{5.8}$$

$$y_t = h(x_t, e_t) \tag{5.9}$$

式中，正如前一节所讨论的，$x \in \mathbb{R}^n$是包括未知参数的状态，并且 $y \in \mathbb{R}^m$是输出。

我们没有把外源输入项 u_t 列入到简单的方程式中，因为 u_t 也是作为问题估计中的一项，从评估的观点来说和 y_t 起了同样的作用。这种简化并不意味着在实验中外援输入是不重要的。从实验设计的角度来看，它明显地被认为是一种重要的变量。

$Y_t = \{y_0, \cdots, y_t\}$，在一般情况下，状态估计被明确地表达为条件分布函数 $p(x_t|Y_t)$的解决方案，也就是说，给出观察的状态的分布决定于当前时间。一旦条件分布函数是已知的，状态估计（点估计）可以计算为

$$\hat{x}_t = \Theta(p(x_t|Y_t))$$

式中，Θ 是一个非线性函数符。

为清楚起见，在结果中我们用 $x_{t|t}$代替 $\hat{x}_t$ 这意味着 x_t 的估计是鉴于所有可用的观测，包括时间。鉴于 $p(x_t|Y_t)$，人们可以根据特定的优化标准计算 x_t 的点值估计。x_t 的平均值估计可以写成

$$\bar{x}_{t|t} = \int x_t p(x_t|Y_t)\,\mathrm{d}x_t$$

最大后验估计（MAP）是

$$\mathrm{MAP}(x_{t|t}) = \mathrm{argmax}\{p(x_t|Y_t)\}$$

图 5.1 举例阐明了平均值和 MAP 估计。

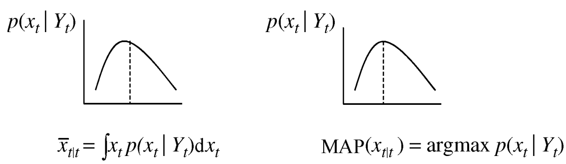


图 5.1　直观点值估计例子

因为分布是已知的，估计值的方差或者其他统计也可以被计算出来。这是整个概率分布函数估计的重要优点。

动态系统状态估计的通解一般由下式给出（Bergman，1999）：

$$p(x_t|Y_{t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|Y_{t-1})\,\mathrm{d}x_t \tag{5.10}$$

$$p(x_t|Y_t) = \frac{p(y_t|x_t)p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \tag{5.11}$$

其中，式（5.10）也被称为 Chapman – Kolmogorov 方程。

式（5.10）和式（5.11）为状态估计提供了一个递归方程。根据状态评估［式（5.8)］，式（5.10）可以计算出来，这个过程称为预测步骤。图 5.2 所示的过程，这里人们可以发现高斯概率分布函数会被曲解，是因为系统的非线性或者 v_t 的非高斯分布。

由于 $p(y_t|x_t)$ 是一个考虑到观察时间 t 的常数，它可以被看作归一化常数。因而，式（5.11）可以写为

$$p(x_t|Y_t) = \alpha p(y_t|x_t)p(x_t|Y_{t-1}) \tag{5.12}$$

式中，$p(y_t|x_t)$ 是 y_t 的概率函数，可以通过式（5.9）计算出来。

这被认为是由新观察值 y_t 得到的更新步骤。这个过程如图 5.3 所示。

总之，序贯贝叶斯推理的状态估计由两个步骤组成：

1）预测步骤：$p(x_t|Y_{t-1})$ 根据状态函数 $f(x_{t-1}, v_{t-1})$ 通过 Chapman－Kolmogorov 方程预测概率分布函数。

2）更新步骤：$p(x_t|Y_{t-1})$ 通过 $p(y_t|x_t)$ 的似然估计乘法更新为 $p(x_t|Y_t)$。

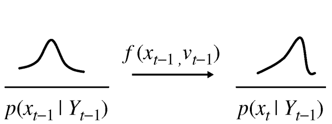


图 5.2　概率分布函数预测

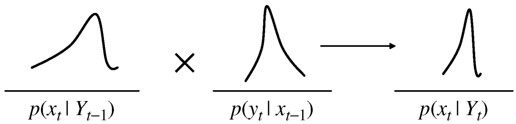


图 5.3　概率分布函数更新

连续的贝叶斯推理过程对状态估计是从 0 时开始的，这从图 5.4 中可以看出。

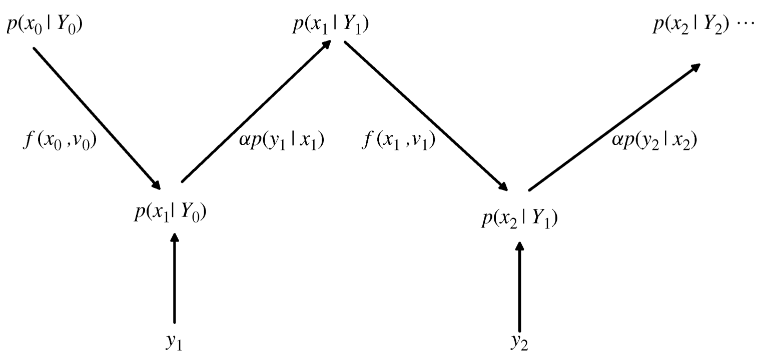


图 5.4　序贯贝叶斯推断法

如果认为 $p(x_t|Y_{t-1})$ 是在前的（也就是在最新的观察 y_t 更新之前），$p(x_t|Y_t)$ 是在后的（也就是更新以后的），这样式（5.12）意味着

$$\text{在后的} \propto \text{可能性} \times \text{在前的} \tag{5.13}$$

这条规则可以应用于任意具有任何分布的非线性系统的过滤。之前的计算也被称为预算，然而在传统的 KF 术语中这种后的计算称为更新。

例 5.3.1　作为应用贝叶斯推理规则的一个例子，考虑一个简单的模型：

$$y = x + e \tag{5.14}$$

这里

$$p(x) = \begin{cases} 0.6 & x = -1 \\ 0.2 & x = 0 \\ 0.2 & x = 1 \end{cases} \tag{5.15}$$

并且

$$p(e) = \begin{cases} 0.2 & e = -1 \\ 0.8 & e = 1 \end{cases} \tag{5.16}$$

可以得出 $y=0$ 是求出的，现在的问题是要找到状态 x 的贝叶斯估计。

根据贝叶斯推理规则，式（5.13），结果可以写成

$$p(x|y=0)=\alpha p(y=0|x)p(x)$$

考虑到式（5.14）和式（5.16），很容易看到 $p\ (y=0|x)$ 可能性的结果：

$p(y=0|x=-1)=0.8$ 也就是说，如果 $y=0$，$x=-1$,那么 e 必须为 1

$p(y=0|x=0)=0$ 也就是说，$y=0$ 与 $x=0$ 是不可能复合的

$p(y=0|x=1)=0.2$ 也就是说，$y=0$，$x=1$，e 必须为 -1

与之前的式（5.15）给出的结果相比，可以得出下面的推论：

$p(x=-1|y=0)=\alpha p(y=0|x=-1)p(x=-1)=\alpha\times0.8\times0.6=0.48\alpha$

$p(x=0|y=0)=\alpha p(y=0|x=0)p(x=0)=\alpha\times0\times0.2=0$

$p(x=1|y=0)=\alpha p(y=0|x=1)p(x=1)=\alpha\times0.2\times0.2=0.04\alpha$

作为概率之和必须为 1，也就是说

$$\begin{aligned}1&=p(x=-1|y=0)+p(x=0|y=0)+p(x=1|y=0)\\&=0.48\alpha+0.04\alpha=0.52\alpha\end{aligned}$$

$$\alpha=1/0.52=1.92$$

状态 x 的贝叶斯推理收益的概率分布函数为

$$p(x|y=0)=\begin{cases}0.92 & x=-1\\0 & x=0\\0.08 & x=1\end{cases}$$

因此，平均值估计和最大后验概率分别可以被计算为

$\bar{x}=-1\times p(x=-1|y=0)+0\times p(x=0|y=0)+1\times p(x=1|y=0)=-0.84$

$\mathrm{MAP}(x)=-1$

这就完成了算例分析。

一般状态估计问题的解决方案会相当复杂。例如，如果部分或者所有的参数都是未知的，即使是一个简单的线性状态空间模型也会转换为一个非线性估计问题。作为一个非线性状态空间模型的特例，一个线性非时变状态空间模型可以通过下式得出：

$$x_{t+1}=Ax_t+v_t \tag{5.17}$$

$$y_t=Cx_t+e_t \tag{5.18}$$

式中，v_t 和 e_t 服从高斯分布。

著名的 KF 在所有的参数是已知的条件下解决了状态估计问题。但是，如果所有或者部分 A 和 C 的参数是未知的，它们作为增强状态必须被估计和明确地表达，例如对于未知参数的 $\theta_{t+1}=\theta_t$。这种线性系统的估计现在变成了非线性状态估计问题。

5.3.1 KF 和 EKF

顺序贝叶斯推理是一个概率分布函数的增值。如果概率分布函数是高斯函数，

概率分布函数中的传播可以被简化为均值矢量和协方差矩阵之间的传播，因为这两个统计数据完全确定高斯概率分布函数。

如果基础系统（underlying system）是线性的噪声污染是高斯函数，KF 是最佳的。即使噪声不是高斯函数，KF 是最佳的决定于在线性估计类别里的最初两个时刻。

作为连续的贝叶斯推理的特例，KF 包括两个步骤：一个预测步骤，紧随其后的是一个更新的步骤。在预测步骤中，先前时间步骤的估计传播到当前的时间步骤。鉴于过去更新估计 $\boldsymbol{x}_{t-1|t-1}$，预测步骤可以用如下的置换解决：

$$\boldsymbol{x}_{t|t-1} = \boldsymbol{E}[f(\boldsymbol{x}_{t-1|t-1}, u_t - 1)]$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{E}[(\boldsymbol{x}_t - \boldsymbol{x}_{t|t-1})(\boldsymbol{x}_t - \boldsymbol{x}_{t|t-1})^{\mathrm{T}}]$$

式中，$\boldsymbol{x}_{t|t-1}$表示 x_t 的预测决定于基于 $t-1$ 所有的观察；$\boldsymbol{P}$ 是状态估计的协方差。

更新步骤可以从下列方程计算：

$$\boldsymbol{x}_{t|t} = \boldsymbol{x}_{t|t-1} + \boldsymbol{L}_t \boldsymbol{\varepsilon}_t$$

$$\boldsymbol{\varepsilon}_t = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}$$

$$\boldsymbol{y}_{t|t-1} = \boldsymbol{E}[h(\boldsymbol{x}_{t|t-1}, \boldsymbol{e}_t)]$$

$$\boldsymbol{P}_{t|t} = \boldsymbol{E}[(\boldsymbol{x}_t - \boldsymbol{x}_{t|t})(\boldsymbol{x}_t - \boldsymbol{x}_{t|t})^{\mathrm{T}}]$$

权重矩阵 $\boldsymbol{L}_t$ 是选择最小化的跟踪更新后的协方差矩阵 $\boldsymbol{P}_{t|t}$，也就是说，来得到最小方差估计，结果可以通过下式得出：

$$\boldsymbol{L}_t = \sum\nolimits_t^{xy} \sum\nolimits_t^{-1}$$

式中

$$\sum\nolimits_t^{xy} = \mathrm{Cov}(\bar{\boldsymbol{x}}_{t|t-1}, \boldsymbol{\varepsilon}_t)$$

$$\sum\nolimits_t = \mathrm{Cov}(\boldsymbol{\varepsilon}_t)$$

式中，$\bar{\boldsymbol{x}}_{t|t-1}$是状态预测误差。因此，更新后的协方差通过下式计算：

$$\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{L}_t \sum\nolimits_t \boldsymbol{L}_t^{\mathrm{T}}$$

因此，如果下述的预期是可用的：预测的状态（预测均值）和它的误差协方差($\boldsymbol{x}_{t|t-1}$，$\boldsymbol{P}_{t|t-1}$)、输出预测误差的协方差、状态预测误差和输出预测误差之间的互协方差，KF 更新过程可以被计算出来。

如果系统是非线性的，这些数量的趋近方法必须应用。在 EKF 中，非线性函数 $f(x_t, u_t)$ 和 $h(x_t, e_t)$ 通过泰勒级数展开线性化，并且二阶和高阶项被忽略。这个过程导致了一阶 EKF。该方法存在线性化误差、对 Jacobian 矩阵的必要性，而且还存在实现的困难。如果非线性化不能被线性化很好地近似，EKF 会产生偏差并且不一致。在这个案例中，高阶近似，二阶或者更高阶可能会被应用。然而，由于对 Hessian 矩阵的需求，一个二阶 EKF 需要广泛的努力实现（Sitz 等人，2002）。

5.3.2 UKF

与趋近非线性函数 $f(x_t, v_t)$ 和 $h(x_t, e_t)$ 不同，UKF 考虑了概率分布函数的近似。

直观地，均值和方差的集合和协方差或者概率分布函数可以通过直接蒙特卡洛模拟估算出来。例如，如果 x 遵循 $p(x)$ 分布，并且对 y 的概率分布函数的确定感兴趣。在第一阶段，x 的 n 个采样根据 $p(x)$ 产生，这样 y 的 n 个采样可以通过映射 $y=f(x)$ 的直接取代计算出来。y 的分布可以通过 y 的映射采样估算出来。经过确定 y 的分布，接下来的步骤是确定 y 的均值和协方差。y 和 x 的互协方差也许可以通过采样数据计算出来。有足够大的 n，均值和方差可以被精确地估计出来。然而，这种方法的计算负荷是巨大的。这种蒙特卡洛方法非常直观，并且可以通过图 5.5 说明，图中星号代表通过 $p(x)$ 得出和通过 $f(x)$ 传播的采样。作为一种选择，蒙特卡洛模拟过程通过可以通过图 5.6 表示，图中星号代表采样，而箭头代表相应的权重。

不是通过随机采样 x 的蒙特卡洛模拟，控制的采样可以被选择用来大大提高工作效率。Julier 和 Uhlmann（2004）和 van der Merwe（2004）介绍的无迹象转换是一种有效的采样方法。控制采样，被称为采样点的最小集合，不是随机得到的；它们被确定地选择，这样它们能表现出特定的特殊属性。结果是，高阶分布的信息可以被认为是固定的小数量的点。然而，确定采样 x 通过 $p(x)$ 被看作与反应分布的特点一样重要。

考虑非线性函数 $y=f(x)$。经过严格的通过扩展 pdf 和近似匹配的一阶和二阶力矩的数学推导项，无迹采样过程可以确定（Julier 和 Uhlmann，2004；van der Merwe，2004）。这个将在下面给出。

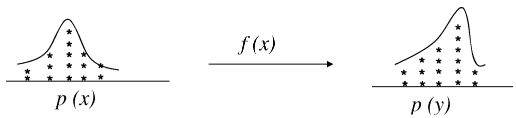


图 5.5 蒙特卡洛估计，这里星号代表根据分布绘制的采样

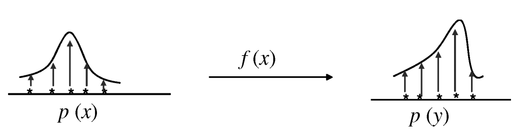


图 5.6 蒙特卡洛估计，这里星号代表采样，箭头代表近似分布的权重

一组 sigma 点包含 $p+1$ 个点，$x^{(i)}$，每个点都有它相应的重量，$w^{(i)}$。提供一组无偏差估计，它们必须遵守下面的条件：

$$\sum_{i=0}^{p} w^{(i)} = 1$$

有了这些点，平均值 $\bar{\boldsymbol{y}}$ 和协方差 Σ_y 可以计算如下：

1）举例说明每个点通过函数得到一组转变 sigma 点：

$$\boldsymbol{y}^{(i)}=f\left(\boldsymbol{x}^{(i)}\right)$$

2）均值通过加权平均的转变点得到：

$$\bar{\boldsymbol{y}}=\sum_{i=0}^{p}\boldsymbol{w}^{(i)}\boldsymbol{y}^{(i)}$$

3）协方差是转变点的加权矢量积：

$$\Sigma_y=\sum_{i=0}^{p}\boldsymbol{w}^{(i)}\left(\boldsymbol{y}^{(i)}-\bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}^{(i)}-\bar{\boldsymbol{y}}\right)^{\mathrm{T}}$$

4）互协方差是原来的点和转变后的点的加权乘积：

$$\Sigma_{xy}=\sum_{i=0}^{p}\boldsymbol{w}^{(i)}\left(\boldsymbol{x}^{(i)}-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{y}^{(i)}-\bar{\boldsymbol{y}}\right)^{\mathrm{T}}$$

满足上述条件的一组 sigma 点包括一组对称的 $2n$ 点，落在$\sqrt{n}$的协方差区域里（Julier 和 Uhlmann，2004）：

$$\boldsymbol{x}^{(i)}=\bar{\boldsymbol{x}}+\left(\sqrt{n\Sigma_x}\right)_i$$

$$\boldsymbol{w}^{(i)}=\frac{1}{2}n$$

$$\boldsymbol{x}^{(i+n)}=\bar{\boldsymbol{x}}+\left(\sqrt{n\Sigma_x}\right)_i$$

$$\boldsymbol{w}^{(i+n)}=\frac{1}{2}n$$

式中，n 是 x 的维数；$\left(\sqrt{n\Sigma_x}\right)_i$ 是 i 行或者 $n\Sigma_x$ 列矩阵的二次方根。无迹转换的两个特点值得注意：（i）无迹转换的计算负荷不大于 EKF，但是没有必要计算 Jacobian 矩阵；（ii）无迹转换计算预测均值和协方差准确到二阶。

例 5.3.2　为了说明上述过程，考虑一个均值/方差估计的 sigma 点的简单例子。

对于具有正态分布的变量 x，两个 sigma 点可以选择为在图 5.7 中所示的 $\mu_x-\sigma_x$ 和 $\mu_x+\sigma_x$，这里 μ_x 和 σ_x 是均值和标准变量的随机方差。如果 $x\sim N$（2，1），这样两个 sigma 点分别是 $x^{(1)}=1$ 和 $x^{(2)}=3$。这两个 sigma 点的相应权重为 1/2。考虑一个映射 $y=5x+1$，这里理论均值和 y 的方差可以被容易地推导出 $\mu_y=11$ 和 $\sigma_y^2=25$。y 和 x 的协方差可以被推导为 $\sigma_{xy}=5$。现在，使用 sigma 点为数值计算，可以得到以下的结果：

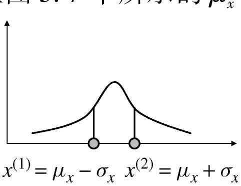


图 5.7　sigma 点说明实例

$$\hat{\mu}_x=1/2x^{(1)}+1/2x^{(2)}=\frac{1}{2}(1+3)=2$$

$$\hat{\sigma}_x^2=1/2(x^{(1)}-\hat{\mu}_x)^2+1/2(x^{(1)}-\hat{\mu}_x)^2=1/2(1-2)^2+1/2(3-2)^2=1$$

这与理论结果完全一致。与这两个 sigma 点一起，相应的映射到 y 分别为

$y^{(1)} = 5x^{(1)} + 1 = 5 \times 1 + 1 = 6$

$y^{(2)} = 5x^{(2)} + 1 = 5 \times 3 + 1 = 16$

以下数据可以通过无迹计算算法计算出来：

$$\hat{\mu}_y = 1/2y^{(1)} + 1/2y^{(2)} = \frac{1}{2}(6 + 16) = 11$$

$$\hat{\sigma}_y^2 = 1/2(y^{(1)} - \hat{\mu}_y)^2 + 1/2(y^{(2)} - \hat{\mu}_y)^2 = 1/2(6 - 11)^2 + 1/2(16 - 11)^2 = 25$$

$$\begin{aligned}\hat{\sigma}_{xy} &= 1/2((x^{(1)} - \hat{\mu}_x)(y^{(1)} - \hat{\mu}_y) + 1/2(x^{(2)} - \hat{\mu}_x)(y^{(2)} - \hat{\mu}_y) \\ &= 1/2\ (1-2)\ (6-11)\ + 1/2\ (3-2)\ (16-11)\ = 5\end{aligned}$$

这些数学结果与理论值完全一致。对于线性映射，这些统计资料可以精确到二阶（Julier 和 Uhlmann，2004；van der Merwe，2004）。

很清楚，这些 sigma 点的选择不是唯一的。Julier 和 Uhlmann（2004）通过引入许多额外的扩展因素，提出了许多 sigma 点的不同选择来继续适应不同的概率分布函数。sigma 点也可能是扩大化的，并且这会对匹配高阶力矩有更多的自由，例如偏差。van der Merwe 实施了更多的无迹采样程序，并表现出 sigma 点的方法与传统加权最小方差的相似性，指出无迹转变的优越性。

小结本节，可以得出 UKF 的下列程序：

1）用式（5.8）和式（5.9）包括 v_t 和 e_t 状态扩充，称为

$$\boldsymbol{X}_t = \begin{bmatrix} x_t \\ v_t \\ e_t \end{bmatrix}$$

新的状态空间方程可以写为

$$\boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t)$$
$$\boldsymbol{y}_t = h(\boldsymbol{X}_t)$$

2）一组 sigma 点的选择可以通过实施一个 sigma 点的选择算法到增强状态得到 sigma 点 $X_{t-1}^{(i)}$，$i = 1$，…，$p + 1$ 和相应的权重 $w^{(i)}$。

3）sigma 点在 $\boldsymbol{X}_t$ 和 $\boldsymbol{y}_t$ 之内的映射通过非线性状态空间方程得出：

$$\boldsymbol{X}_t^{(i)} = f(\boldsymbol{X}_{t-1}^{(i)})$$
$$\boldsymbol{y}_t^{(i)} = h(\boldsymbol{X}_{t-1}^{(i)})$$

4）预测状态和输出的均值可以被计算为

$$\boldsymbol{X}_{t|t-1} = \sum_{i=0}^{p} \boldsymbol{w}^{(i)} \boldsymbol{X}_t^{(i)}$$

$$\boldsymbol{y}_{t|t-1} = \sum_{i=0}^{p} \boldsymbol{w}^{(i)} \boldsymbol{y}_t^{(i)}$$

5）方差和协方差可以计算为

$$\boldsymbol{P}_{t|t-1} = \sum_{i=0}^{p} \boldsymbol{w}^{(i)} (\boldsymbol{X}_t^{(i)} - \boldsymbol{X}_{t|t-1})(\boldsymbol{X}_t^{(i)} - \boldsymbol{X}_{t|t-1})^{\mathrm{T}}$$

$$\Sigma_t = \sum_{i=0}^{p} \boldsymbol{w}^{(i)} \left(\boldsymbol{y}_t^{(i)} - \boldsymbol{y}_{t|t-1}\right)\left(\boldsymbol{y}_t^{(i)} - \boldsymbol{y}_{t|t-1}\right)^{\mathrm{T}}$$

$$\Sigma_t^{xy} = \sum_{i=0}^{p} \boldsymbol{w}^{(i)} \left(\boldsymbol{X}_t^{(i)} - \boldsymbol{X}_{t|t-1}\right)\left(\boldsymbol{y}_t^{(i)} - \boldsymbol{y}_{t|t-1}\right)^{\mathrm{T}}$$

6）更新步骤可以通过下式得出：

$$\boldsymbol{X}_{t|t} = \boldsymbol{X}_{t|t-1} + \boldsymbol{L}_t \boldsymbol{\varepsilon}_t$$

$$\boldsymbol{\varepsilon}_t = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}$$

$$\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{L}_t \Sigma_t \boldsymbol{L}_t^{\mathrm{T}}$$

式中

$$\boldsymbol{L}_t = \Sigma_t^{xy} \Sigma_t^{-1}$$

在图 5.8 中，类比蒙特卡洛模拟，UKF 的过程为一个标量状态推理。本质上是 UKF 的“几个” sigma 点来近似整个概率分布函数的传播。UKF 和 EKF 的主要区别是近似方法。EKF 接近非线性映射，UKF 接近概率分布函数。

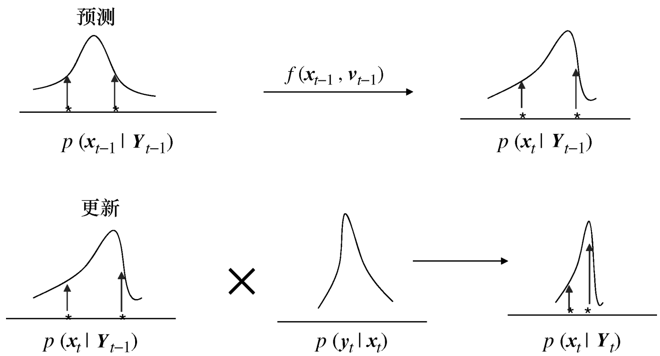


图 5.8　UKF 的推理过程实例，这里星号代表 sigma 点，箭头代表相应的权重

5.4　范例

使用 UKF 的几个实例的状态和参数估计在本节叙述。

例 5.4.1　考虑一个二阶放热反应（Muske 和 Edgar，1997）：

$$\frac{\mathrm{d}C_A}{\mathrm{d}t} = -k_0 \exp(-E_a/RT)\, C_A^2$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{\Delta H}{\rho C} k_0 \exp(-E_a/RT)\, C_A^2 + \frac{UA}{V\rho C}(T_c - T)$$

式中，T_c 是一个被控变量，相关的参数有

$$\mathrm{par1} = -\frac{\Delta H}{\rho C} k_0 = -30$$

$$\mathrm{par2} = \frac{UA}{V\rho C} = 0.001$$

温度 T 是基于测量的，现在的问题是如何估计状态和两个参数。

一组 600 个数据点是通过采样时间 0.05s 仿真的。T_c 是一个随机输入序列。初始浓度为 5mol/L，初始温度为 20℃。应用 UKF，状态和参数的初始猜测为 C_0 = 2mol/L、T_0 = 30℃、par1 = −20 和 par2 = 0.5。估计的结果在图 5.9 中表示出来。

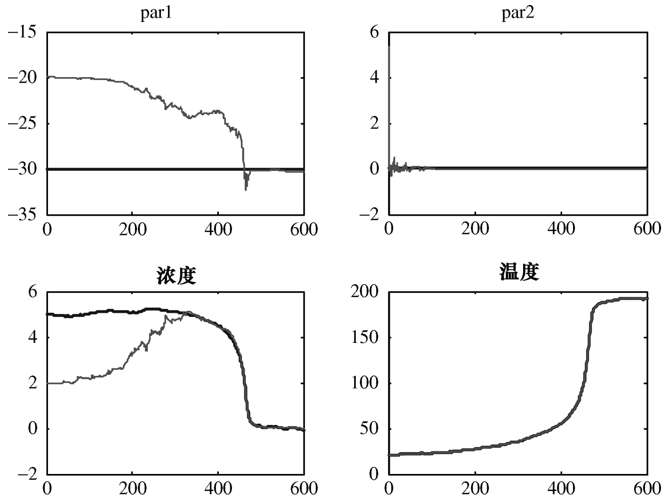


图 5.9　二阶放热反应

这里，所有状态和参数的收敛是可观察到的。UKF 已经被应用到上述模型的估计，它是高度非线性的，但是结果不收敛。原因可能是系统参数的敏感性（参数的低识别性）或者 UKF 很强的非线性的存在的限制。

例 5.4.2　这是 Gordon 等人（1993）讨论的一个已知的高度非线性系统：

$$x_t = 0.5x_{t-1} + \frac{25x_{t-1}}{1 + x_{t-1}^2} + u_{t-1} + v_{t-1}$$

$$y_t = \frac{x_t^2}{20} + e_t$$

$$u_t = 8\cos(1.2(t-1))$$

$$v_t \sim N(0,\ 10)$$

$$e_t \sim N(1,\ 1)$$

这里相关的 3 个参数是上述方程中的 3 个系数：

$$\text{par1} = 0.5$$

$$\text{par2} = 25$$

$$\text{par3} = 20$$

一组 2000 个数据点与 $u_t = 8\cos(1.2(t-1))$ 和最初状态 $x_0 = 0.1$ 一起进行仿真。应用 UKF，状态和参数最初的猜测选择为 $x_0 = 0.15$、par1 = 1、par2 = 20 和 par3 = 25。状态和 3 个参数的估计结果在图 5.10 中给出，其中的状态与前两个参

数的收敛被观测到，而第 3 个参数有一个偏置误差。

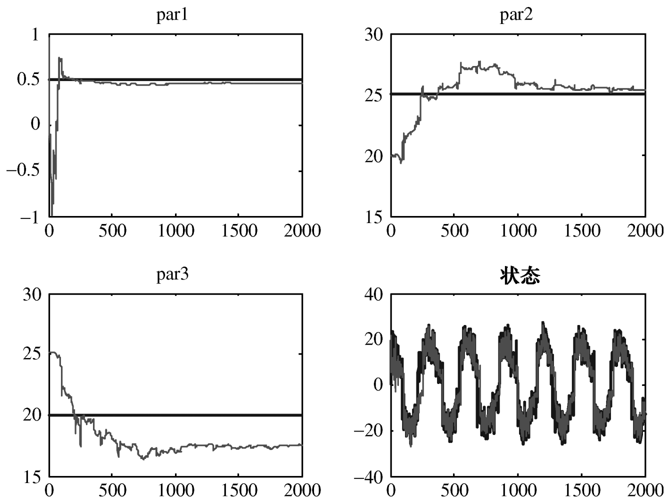


图 5. 10　高度非线性系统的参数和状态估计

例 5. 4. 3　Nowak（2001）提出了病毒动态模型。动态模型由下式给出：

$$\dot{x}(t) = \lambda - dx(t) - \beta x(t)v(t) + e_1(t)$$

$$\dot{y}(t) = \beta x(t)v(t) - ay(t) + e_2(t)$$

$$\dot{v}(t) = ky(t) - u(t)v(t) + e_3(t)$$

式中，$e_1(t)$、$e_2(t)$ 和 $e_3(t)$ 是动态分布 N（0，10^{12}）的干扰；$x(t)$ 代表健康细胞的数量；$y(t)$ 是感染细胞的数量；v 表示许多游离的病毒，如图 5. 11 所示。$y(t)$ 和 $v(t)$ 都测量了，相关的参数为

$$k = 100$$

$$u = 5$$

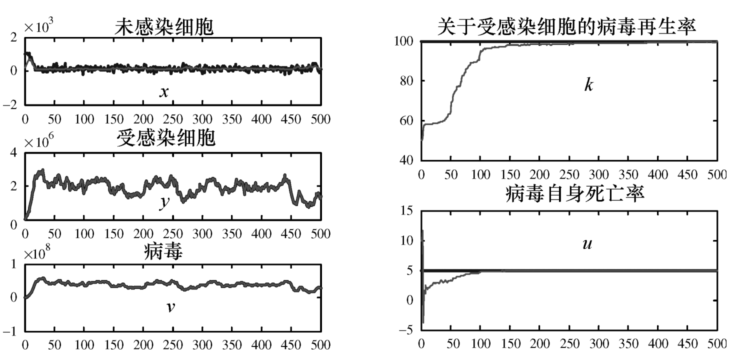


图 5. 11　细胞动态模型的参数和状态估计

一组 500 个数据点在初始状态 $x(0)=10^5$、$y(0)=10^2$ 和 $v(0)=10^3$ 下仿真。采样时间为 0.1s。状态和参数的最初猜测选择为 $x(0)=10^4$、$y(0)=0$、$v(0)=0$、$k=50$ 和 $u=10$。估计的结果如图 5.11 所示，其中状态和参数都收敛到真值。

例 5.4.4 下面的模型是由细胞毒性过程派生的机理模型（Huang 和 Xing，2006）：

$$\frac{\mathrm{d}c_i}{\mathrm{d}t}=k_3\left(k_1c_e+\frac{k_2c_e}{K_i+c_e}-c_i\right)+v_1$$

$$\frac{\mathrm{d}N}{\mathrm{d}t}=(a_0+a_1c_i+a_2c_e)N+v_2$$

式中，v_1 和 v_2 是具有高斯分布 $N(0,\ 0.01)$ 的干扰；c_i 是细胞内的毒物浓度；N 是细胞群体。

被测变量是具有测量噪声 $e\sim N(0,\ 0.01)$ 的细胞群体。相关参数为细胞增殖速率 a_0、细胞凋亡率 a_1 和坏死系数 a_2，下面为真实数值：

$$a_0=0.031$$

$$a_1=1$$

$$a_2=-0.381$$

一组 1000 个数据点在随机多层外部扰动在细胞外浓度 c_e 下模拟。初始状态为 $c_i(0)=1$ 和 $N(0)=1$。离散时间为 $\Delta t=1/3600\text{h}=1\text{s}$。最初的状态猜测和参数选择为 $c_i(0)=0.5$、$N(0)=0.3$、$a_0=0$、$a_1=0$ 和 $a_2=0$。结果估计如图 5.12 所示，所有状态和参数收敛到真值。

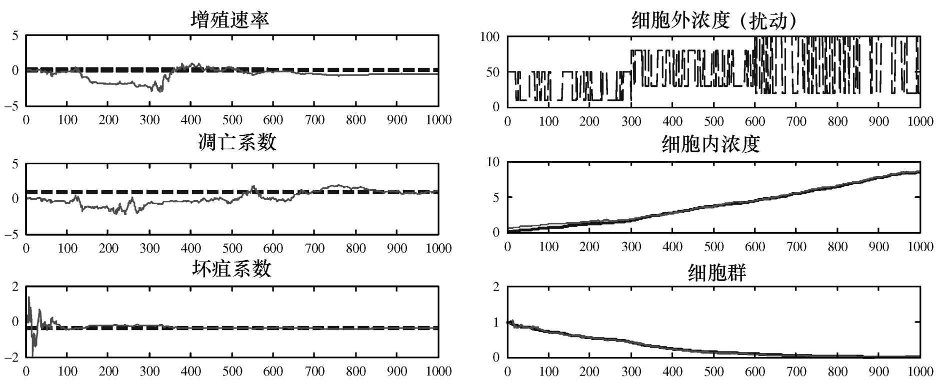


图 5.12 细胞毒性动态模型参数和状态估计

5.5 注解和参考文献

本章集中于 UKF，它会应用于本书中的 SOFC 的状态估计。它基于第一作者的

陈述（Huang 和 Wang，2006），并且主要参考文献作者为 Gustafsson 和 Hriljac（2003）、Julier 和 Uhlmann（1997 和 2004）、Norgaard 等人（2000）、Sitz 等人（2000）、van der Merwe（2004）、Wan 和 van der Merwe（2000 和 2001）、Wan 等人（2000）。对于 EnKF 的介绍，参考 Evensen（1994 和 2009）的研究。对于对粒子滤波器感兴趣的读者，参考 Doucet 等人（2001）、Gordon 等人（1993）、van der Merwe（2004）的研究进一步阅读。一个非线性状态评估算法的详细概述，与它们的一些实际应用可以在 Ristic 等人（2004）的文献中找到。

第 6 章　模型预测控制

6.1　最先进的模型预测控制

在过去数十年中，线性模型预测控制（MPC）及其变型已成为使用最广泛的先进控制技术。在许多情况下，先进的控制就是指 MPC。大多数工业应用中使用线性模型和线性约束。通过在最佳条件下控制设备，可以体现 MPC 的优势。在大多数情况下，最佳条件在边界条件之外或就在边界上，因此，通过在约束条件上控制设备能够实现最大收益。由于大多数流程在本质上是非线性的，采用线性模型和约束条件可能不会找到最大收益点。在这样的情况下，线性模型无法提供正确的动态系统，因此，设计的 MPC 需要非线性模型。

MPC 中，一系列未来控制指令，其中包括当前的控制指令，都是基于系统模型上进行计算的。所以称它为基于模型的预测控制。该模型可以是线性或非线性的。其主要目的是尽量简化对于模型方程和一些物理限制的目标函数（通常是一个关于状态和输入的二次函数）。对于线性时不变模型，这个问题是可以通过给出函数的输入和输出作为控制指令来解决的。对于非线性的情况，对于优化问题，通常没有明确的解决方案，只能在数值上被迫来解决它。

在过去的几十年中，线性 MPC 已经成功地应用到不同的行业中来解决问题，特别是化学和碳氢化合物行业。应用范围包括锅炉涡轮调节、氮气纯化柱、蒸汽温度控制、蒸馏塔控制、燃烧控制和优化等。

另一方面，非线性对应的 MPC，由于其固有的复杂性所以并没有达到同样的成功。除了复杂的非线性模型的要求外，非线性 MPC 应用的主要缺点就是大量的计算负荷。一些研究人员正在为发展高效和强大的非线性 MPC 努力着，特别是其在工业应用中。

最近，人们对非线性 MPC 的兴趣持续增长。这是因为，今天的工艺过程中需要在制约附近进行操作，以改善企业的盈利能力，以满足更严格的环保和安全法规要求。如果固有的非线性过程近似替换线性模型，那么满足这些要求更严格，广义的延伸 MPC 的非线性 MPC 非常适合于这种情况。

MPC 的主要特点和特性可以归纳如下：

模型。即使机理模型需要被离散化，以适应 MPC 结构，但它允许直接利用模型进行预测。

约束处理能力。MPC 的最大的优势或许是，它可以处理输入、输出和状态约束。

输出预测。其进行在线优化，找出未来的控制指令与可预测的输出来实现控制目标，因其在预测时域（PH）中进行，所以故名预测控制。

滚动时域反馈控制。只实施第一个计算控制命令，在线优化就会在每个采样间隔重复。为了更新基于模型的预测，测量在每个实例中都得到使用，这对控制律提供了连续的反馈。

状态估计。在状态空间公式化下，状态可以被控制，但并非所有状态都可以直接测量。对于不可测状态需要估计。

6.2　基本原理

MPC 可以被描述为，考虑到约束条件，有限域中的开环最优控制问题。在当前时间对状态测量或估计的基础上，可以预测未来状态和所需的控制指令，使得在预期域中定义的目标函数最小化。然后，它应用第一个计算控制指令，然后进入到下一个时间步长，并重复整个过程。由于 MPC 目标函数是使用的当前状态，所以该解决方案实际上利用的是在每一个时间步长下的反馈数据。因此，MPC 算法收敛于一个封闭的回路策略。MPC 的滚动时域功能，区别于原来的控制器，其如图 6.1 所示。

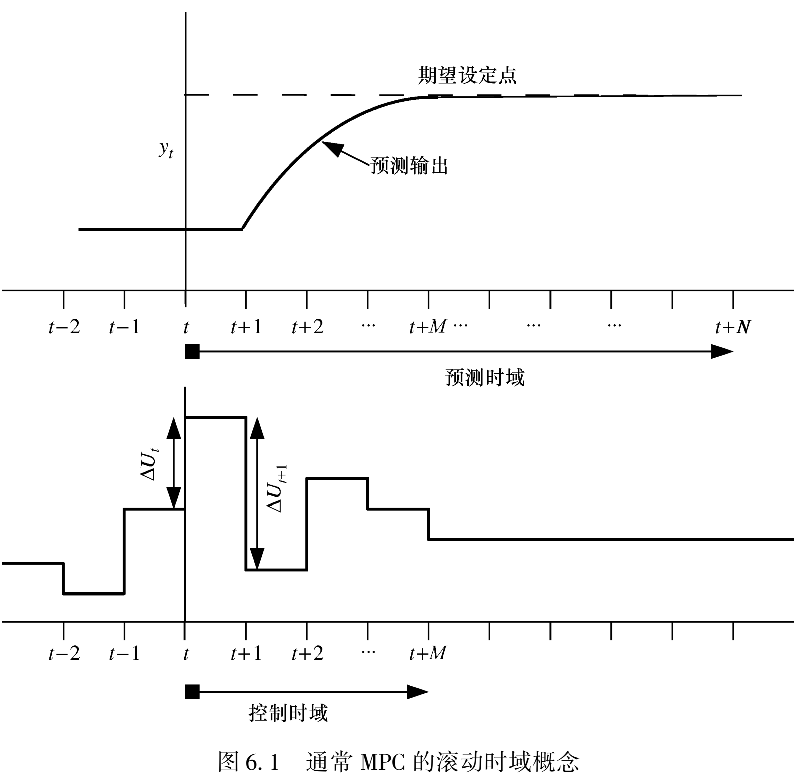


图 6.1　通常 MPC 的滚动时域概念

6.2.1 MPC 模型

不同策略的 MPC，可以使用多种方法来表示输出和可测输入的关系，其中一些是操纵变量被认为是可测的扰动。为了描述没有被工艺模型反应的行为，需要考虑干扰模型，包括不可测量的输入、噪声和模型误差的影响。因此，该模型可以被分成两部分：实际过程模型和扰动模型。以下是 MPC 常用模型的例子。

脉冲响应模型由下式得出：

$$y_t = \sum_{k=1}^{\infty} h_k u_{t-k}$$

式中，h_k在单一脉冲扰动下的过程的采样输出。

其总和被截断为一个稳定的过程，通常是在其稳定时间 N_s内，并且只考虑 N_s值。所以有

$$y_t = \sum_{k=1}^{N_s} h_k u_{t-k} = H(z^{-1})u_t$$

式中，$H(z^{-1}) = h_1 z^{-1} + h_2 z^{-2} + \cdots + h_{N_s} z^{-N_s}$。

脉冲响应模型是直观的，并且识别需要无先前的信息（如有关过程模型的顺序）。然而，没有积分器的稳定的过程可以由一个有限脉冲响应的模型表示。此外，还需要大量的参数。

或者，一阶响应模型可以被认为是

$$y_t = \sum_{k=1}^{\infty} g_k \Delta u_{t-k} = G(z^{-1})(1 - z^{-1})u_t \tag{6.1}$$

式中，g_k是对于单位阶跃输入而言的采样输出值；$\Delta u_t = u_t - u_{t-1}$。在阶跃响应系数中脉冲不同，因此得到

$$h_k = g_k - g_{k-1}$$

$$g_k = \sum_{j=1}^{k} h_j$$

阶跃响应模型常用做 PID 控制调节，并且是在 MPC 设计中最流行的。

另一种常用的模型是传递函数模型，传递函数模型由下式给出：

$$y_t = \frac{B(z^{-1})}{A(z^{-1})}u_t$$

它是有效的不稳定的过程，需要更少的参数来描述模型。然而，由于在模型中的参数的数量有限，它有一定捕获复杂动态过程的局限性。此外，该模型的阶必须在要估计的模型参数之前确定。

另一种流行的模式是状态空间模型，状态空间模型如

$$x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t$$

其通常用于模拟线性多变量过程。然而，它可能会因为未知的测量状态而复杂化，从而需要进行额外增加观察。

扰动模型原则上可以采取与上面讨论的过程模型相同的形式，其中输入的 u_t 被替换为噪声 e_t。然而，最常用的扰动模型是随机变动的，其表示为

$$v_t = \frac{1}{1 - z^{-1}} e_t$$

式中，v_t 是过程扰动影响；e_t 是白噪声。

当用状态空间模型来表示过程时，动态扰动便自然地融入到状态空间模型。例如，创新模式：

$$x_{t+1} = Ax_t + Bu_t + Ke_t$$
$$y_t = Cx_t + e_t$$

它们都包含了过程和动态扰动。

6.2.2　自由和强迫响应

设计 MPC 时，常用自由和强迫响应的概念。这里的想法是将过程响应变为两个信号的总和：

$$y_t = y_f(t) + y_c(t)$$

如果操作变量在通过 PH 值时，保持恒定自由响应的，则 $y_f(t)$ 对应于输出的预测。另一方面，由于未来的增量控制，强制响应，$y_c(t)$ 对应于输出的预测。

为了理解概念，控制序列可表示为两个项的和：

$$u_t = u_f(t) + u_c(t)$$

在未来时刻，信号 $u_f(t)$ 保持恒定且等于操纵变量的最后一个值，也就是说：

$$u_f(t-j) = u_{t-j} \quad j = 1, 2, \cdots$$
$$u_f(t+j) = u_{t-1} \quad j = 0, 1, 2, \cdots$$

信号的 $u_c(t)$ 在过去等于零，在未来等于增量控制移动：

$$u_c(t-j) = 0 \quad j = 1, 2, \cdots$$
$$u_c(t+j) = u_{t+j} - u_{t-1} \quad j = 0, 1, 2, \cdots$$

这样，自由响应源于 u_f，因为其目前的状态，它对应于过程的进展。强迫响应源于 u_c，并且它的响应与未来控制移动相关。

6.2.3　目标函数

各种 MPC 算法提出了不同的成本函数。总的目标是（Camacho 和 Bordons，1999）：

- 未来的输出应该在考虑视野上遵循一个确定的参考信号；
- 在目标函数中，这样做的控制工作的必要性应加以考虑。

这样的目标函数的例子是

$$J=\sum_{j=N_1}^{N_2}\left[y_{t+j}^{\mathrm{ref}}-\hat{y}(t+j|t)\right]^{\mathrm{T}}\boldsymbol{Q}\left[y_{t+j}^{\mathrm{ref}}-\hat{y}(t+j|t)\right] + \sum_{j=1}^{M}\left[\Delta u_{t+j-1}\right]^{\mathrm{T}}\boldsymbol{R}\left[\Delta u_{t+j-1}\right] \tag{6.2}$$

当预测和控制的视野设定为相同且 $N_1=1$ 时，目标函数可写为

$$J=\sum_{j=1}^{N_2}\left[\left\|y_{t+j}^{\mathrm{ref}}-\hat{y}(t+j|t)\right\|_{\boldsymbol{Q}}^2+\left\|\Delta u_{t+j-1}\right\|_{\boldsymbol{R}}^2\right]$$

式中，Q 和 R 分别是加权矩阵；y_t^{ref} 为输出或设置点的参考信号；$\hat{y}(t+j|t)$ 是在时刻 t 根据所有可用的历史输出和所有输入的值 y_{t+j} 预测；$\Delta u_t=u_t-u_{t-1}$ 是增量控制动作或转换速率。

预测开始于 N_1 或 1 时，N_2 是最大的 PH 值，M 是控制时域。N_2-N_1+1 决定了预测窗口，其中根据设定点预测输出是理想的。因此，采用 N_1 的一个较大的值意味着如果在最初的几个瞬间有错误使其达到 N_1，它也不会有太大影响。然而，采用 N_2-N_1+1 的较大值，意味着需要在一个较长的时间跨度中关注输出误差。

6.2.4　限制条件

在实践中，所有的进程都受到限制。执行机构具有有限的动作领域以及有限的转换速率，如在有阀的情况下，受限于完全打开或关闭的位置以及应答率。推测其原因，是安全或环境，甚至在传感器、示波器本身都可引起的过程变量中的限制，例如在罐中的水平、管中的流动、温度或压力。运作上的限制由市场而定。例如，如果对于某个产品没有需求，最大约束可能需要将该产品放入高级过程控制（APC）中。所有这些使得在函数中引入的限制能被最小化。通常情况下，振幅的边界、在控制信号转化率的边界和输出的边界都要考虑到，如下所示：

$$u_{\min}\leqslant u_t\leqslant u_{\max}$$
$$\Delta u_{\min}\leqslant\Delta u_t\leqslant\Delta u_{\max}$$
$$y_{\min}\leqslant y_t\leqslant y_{\max}$$

6.2.5　MPC 规则

为了获得受控制的输入，u_{t+p} 必须使目标函数最小化。要做到这一点，$\hat{y}(t+j|t)$ 的预测输出值被当作过去的输入、输出的函数来计算。未来的控制信号使用所选择的模型，然后在成本函数中替换它们。它们的最小化给出寻找值的表达式。如果该模型是线性的，并且没有限制，分析方案可以通过二次判据来获得，否则应采用数值方法。

对于任何方法，得到解决是不容易的，因为有许多变量，其数量可以很高。为了减少自由度，一定的结构可能施加在控制规则上。此外，已经发现，这种结构的控制规则提高了系统的鲁棒性和总体表现，基本上是因为允许操纵

变量的自由演化可能导致不愿得到的高频控制信号处于最差的、不稳定的状态。

所述控制时域可用于在控制规则上施加一个结构。这样的理念，可以认为，在一定间隔 $M < N_2$ 后，该信号变为恒定，即

$$\Delta u_{t+j-1} = 0 \qquad j > M$$

6.3 DMC

动态矩阵控制（DMC）最初是由壳牌石油公司的 Cutler 和 Ramaker 开发的，它是最有名的、最实用的 MPC 算法。本节将简要介绍 DMC，包括预测的发展、控制算法的推导和扰动处理的推导。

6.3.1 预测

从工业过程起源来讲，其中阶跃响应模型通常用于调整控制器，DMC 还采用了阶跃响应模型来描述过程动态。阶跃响应模型既可以被转换成其他参数模型，也可以直接从阶跃响应测试估计。阶跃响应模型已引入式（6.1），其可以进一步分离成两项：

$$y_t = \sum_{k=1}^{\infty} g_k \Delta u_{t-k} = \sum_{k=1}^{N_s} g_k \Delta u_{t-k} + \sum_{k=N_s+1}^{\infty} g_k \Delta u_{t-k} \tag{6.3}$$

式中，N_s 被定义为设定时间或有限脉冲响应模型的截断点。

式（6.3）的第一项是有限的阶跃响应模型。第二项表示累计响应，因为控制行为是从无限的过去一直到时间点 $t - N_s - 1$。用 Z_t 表示：

$$Z_t = \sum_{k=N_s+1}^{\infty} g_k \Delta u_{t-k} \tag{6.4}$$

式（6.3）可写为

$$y_t = \sum_{k=1}^{N_s} g_k \Delta u_{t-k} + Z_t \tag{6.5}$$

在式（6.5）中通过 p 采样来改变时间参数得到：

$$y_{t+p} = \sum_{k=1}^{N_s} g_k \Delta u_{t+p-k} + Z_{t+p} \tag{6.6}$$

式（6.6）可进一步扩展：

$$y_{t+p} = \sum_{k=1}^{p} g_k \Delta u_{t+p-k} + \sum_{k=p+1}^{N_s} g_k \Delta u_{t+p-k} + Z_{t+p} \tag{6.7}$$

这里基准时间设定为 t，式（6.7）的第一项包含未来的控制动作（Δu_t，Δu_{t+1}，…），其被称为强迫响应，第二项和第三项包含过去的控制动作（Δu_{t-1}，Δu_{t-2}，…），其为已知的自由响应。自由响应表示为

$$y_{t+p}^{*} = \sum_{k=p+1}^{N_s} g_k \Delta u_{t+p-k} + Z_{t+p} \tag{6.8}$$

式（6.7）可进一步写成

$$y_{t+p} = \sum_{k=1}^{p} g_k \Delta u_{t+p-k} + y_{t+p}^{*} \tag{6.9}$$

正如前面已经在 MPC 的设计框架下讨论的，MPC 的预测响应由两部分组成：自由响应项和强迫响应。因此，式（6.9）是对 MPC 的预测。作为相对于时间点 t 的第一项由未来控制动作组成，式（6.9）提供超前 y 预测的 p 阶。为了反映预测性质，用 $\hat{y}$（$t+p \mid t$）替换 $y_{t+p}p$ 阶超前预测可以写成

$$\hat{y}(t+p \mid t) = \sum_{k=1}^{p} g_k \Delta u_{t+p-k} + y_{t+p}^{*} \tag{6.10}$$

MPC 目标函数表明，用式（6.2）设计 MPC，对于在有限的时域 y 值的预测是必要的。使用式（6.10），从一阶到 N 阶的预测［即相当于在式（6.2）中 $N_1=1$ 和 $N_2=N$］可以得出

$$\begin{aligned}
\hat{y}(t+1 \mid t) &= g_1 \Delta u_t + y_{t+1}^{*} \\
\hat{y}(t+2 \mid t) &= g_2 \Delta u_t + g_1 \Delta u_{t+1} + y_{t+2}^{*} \\
&\vdots \\
\hat{y}(t+M \mid t) &= g_M \Delta u_t + g_{M-1} \Delta u_{t+1} + \cdots + g_1 \Delta u_{t+M-1} + y_{t+M}^{*} \\
&\vdots \\
\hat{y}(t+N \mid t) &= g_N \Delta u_t + g_{N-1} \Delta u_{t+1} + \cdots + g_1 \Delta u_{t+N-1} + y_{t+N}^{*}
\end{aligned}$$

需要注意的是，当控制时域设置为 M 时，$\Delta u_{t+M} = \Delta u_{t+M+1} = \cdots = \Delta u_{t+N-1} = 0$。

这些预测可以写在一个更紧凑的矢量/矩阵方程中：

$$\hat{\boldsymbol{y}} = \boldsymbol{y}^{*} + \boldsymbol{G}\Delta\boldsymbol{u} \tag{6.11}$$

式中，$\boldsymbol{G}$ 为

$$G = \begin{bmatrix} g_1 & 0 & 0 & \cdots & 0 \\ g_2 & g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_M & g_{M-1} & g_{M-2} & \cdots & g_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_N & g_{N-1} & g_{N-2} & \cdots & g_{N-M+1} \end{bmatrix}$$

其被称为动态矩阵，另外：

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \hat{y}(t+1 \mid t) \\ \hat{y}(t+2 \mid t) \\ \vdots \\ \hat{y}(t+N \mid t) \end{bmatrix}$$

$$\boldsymbol{y}^* = \begin{bmatrix} y_{t+1}^* \\ y_{t+2}^* \\ \vdots \\ y_{t+N}^* \end{bmatrix}$$

$$\Delta\boldsymbol{u} = \begin{bmatrix} \Delta u_t \\ \Delta u_{t+1} \\ \vdots \\ \Delta u_{t+M-1} \end{bmatrix}$$

6.3.2 无变形控制移动的 DMC

根据预测的情况，准备推导出 DMC 算法。首先，考虑一个 MPC 算法，其仅减少了由式（6.2）所定义的 MPC 目标函数的第一项。固定点轨迹 PH 值通过以下确定：

$$\boldsymbol{y}^{\mathrm{ref}} = \begin{bmatrix} y_{t+1}^{\mathrm{ref}} \\ y_{t+2}^{\mathrm{ref}} \\ \vdots \\ y_{t+N}^{\mathrm{ref}} \end{bmatrix}$$

然后，DMC 算法可以通过下列优化问题导出：

$$\underset{\Delta\boldsymbol{u}}{\text{minimise}} \quad J(\Delta\boldsymbol{u}) = (\boldsymbol{y}^{\mathrm{ref}} - \hat{\boldsymbol{y}})^{\mathrm{T}} \mathscr{Q} (\boldsymbol{y}^{\mathrm{ref}} - \hat{\boldsymbol{y}})$$

限制条件：$\hat{\boldsymbol{y}} = \boldsymbol{y}^* + \boldsymbol{G}\Delta\boldsymbol{u}$

式中，$\mathscr{Q}$是通过 PH 的输出值的加权矩阵，它通常是一个块对角矩阵，其由式（6.2）中定义的各个权重矩阵 $\boldsymbol{Q}$ 组成。

目标函数中取代 $\hat{\boldsymbol{y}}$ 的表达式得到：

$$\underset{\Delta\boldsymbol{u}}{\text{minimise}} \quad J(\Delta\boldsymbol{u}) = (\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^* - \boldsymbol{G}\Delta\boldsymbol{u})^{\mathrm{T}} \mathscr{Q} (\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^* - \boldsymbol{G}\Delta\boldsymbol{u})$$

求 $J(\Delta\boldsymbol{u})$ 对 $\Delta\boldsymbol{u}$ 的偏导然后将其归 0 得

$$\frac{\partial J}{\partial \Delta\boldsymbol{u}} = -2\boldsymbol{G}^{\mathrm{T}} \mathscr{Q}^{\mathrm{T}} (\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^* - \boldsymbol{G}\Delta\boldsymbol{u}) = 0$$

控制动作（也称为压摆率）的最优序列可以从下式得到：

$$\Delta\boldsymbol{u} = (\boldsymbol{G}^{\mathrm{T}} \mathscr{Q}^{\mathrm{T}} \boldsymbol{G})^{-1} \boldsymbol{G}^{\mathrm{T}} \mathscr{Q}^{\mathrm{T}} (\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^*)$$

6.3.3 变形控制移动的 DMC

当考虑式（6.2）中 MPC 的目标函数的第二项时，控制转移将被抑制或惩罚来实现由式（6.2）中第一项测量的控制性能和第二项测量的控制成本之间的转换。所以需要解决下面的优化问题：

$$\underset{\Delta\boldsymbol{u}}{\text{minimise}} \quad J(\Delta\boldsymbol{u}) = (\boldsymbol{y}^{\mathrm{ref}} - \hat{\boldsymbol{y}})^{\mathrm{T}} \mathscr{Q} (\boldsymbol{y}^{\mathrm{ref}} - \hat{\boldsymbol{y}}) + \Delta\boldsymbol{u}^{\mathrm{T}} \mathscr{R} \Delta\boldsymbol{u}$$

限制条件：

$$\hat{\boldsymbol{y}} = \boldsymbol{y}^* + \boldsymbol{G}\Delta\boldsymbol{u}$$

式中，$\mathscr{R}$ 和 $\mathscr{Q}$ 的定义一样，定义为控制动作的块加权矩阵。

在目标函数中代替 $\hat{\boldsymbol{y}}$ 的表达。求 J（Δu）对 Δu 的偏导然后将其归 0 得

$$\frac{\partial J}{\partial \Delta \boldsymbol{u}} = -2\boldsymbol{G}^{\mathrm{T}}\mathscr{Q}^{\mathrm{T}}(\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^* - \boldsymbol{G}\Delta\boldsymbol{u}) + 2\mathscr{R}^{\mathrm{T}}\Delta\boldsymbol{u} = 0$$

最优控制算法为

$$\Delta\boldsymbol{u} = (\boldsymbol{G}^{\mathrm{T}}\mathscr{Q}^{\mathrm{T}}\boldsymbol{G} + \mathscr{R})^{-1}\boldsymbol{G}^{\mathrm{T}}\mathscr{Q}^{\mathrm{T}}(\boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{y}^*) \tag{6.12}$$

式中，$\mathscr{R}$ 反映了控制动作的权重和花费。我们可以看到控制成本越高，控制动作的针对越少。

6.3.4　DMC 算法反馈

除了上述预测功能外，反馈校正也是必要的，其能起到补偿干扰的作用，这通常由包括控制规则中的当前程序输出。在 DMC 中，常假定存在干扰，但在沿着整个 PH 中当前值保持恒定，即如果该干扰是 v_t 在当前时间 t 下，然后在 $t+p$ 下，如果 p 与 PH 相联系，则 $v_{t+p} = v_t$。

当考虑扰动时，自由响应由式（6.8）表示，其包括扰动 v_{t+p}：

$$y_{t+p}^* = Z_{t+p} + \sum_{k=p+1}^{N_s} g_k \Delta u_{t+p-k} + v_{t+p} \tag{6.13}$$

由于扰动，预测的输出不再与实际输出相同，并且差额是由扰动造成的。根据定义，y 的超前一阶预测可以直接从阶跃响应模型推导：

$$\hat{y}(t \mid t-1) = \sum_{k=1}^{\infty} g_k \Delta u_{t-k}$$

因此，扰动可由下式决定：

$$v_{t+p} = v_t = y_t - \hat{y}(t \mid t-1) = y_t - \sum_{k=1}^{\infty} g_k \Delta u_{t-k} \tag{6.14}$$

根据式（6.4），可以得出：

$$Z_{t+p} = \sum_{i=N_s+1}^{\infty} g_k \Delta u_{t+p-k} \tag{6.15}$$

由式（6.15）、式（6.14）和式（6.13）得

$$\begin{aligned} y_{t+p}^* &= \sum_{i=N_s+1}^{\infty} g_k \Delta u_{t+p-k} + \sum_{k=p+1}^{N_s} g_k \Delta u_{t+p-k} + y_t - \sum_{k=1}^{\infty} g_k \Delta u_{t-k} \\ &= y_t + \sum_{k=p+1}^{N_s} g_k \Delta u_{t+p-k} + \sum_{k=N_s+1}^{\infty} g_k \Delta u_{t+p-k} - \sum_{k=1}^{\infty} g_k \Delta u_{t-k} \\ &= y_t + \sum_{k=p+1}^{\infty} g_k \Delta u_{t+p-k} - \sum_{k=1}^{\infty} g_k \Delta u_{t-k} \end{aligned}$$

用 $j = k - p$ 代换变量中的第二项得

$$y^*_{t+p}=y_t+\sum_{j=1}^{\infty}g_{j+p}\Delta u_{t-j}-\sum_{k=1}^{\infty}g_k\Delta u_{t-k} \tag{6.16}$$

式（6.16）中变量从 j 变化到 k 得

$$y^*_{t+p}=y_t-\sum_{k=1}^{\infty}(g_k-g_{k+p})\Delta u_{t-k}$$

N_s 为设定时间得

$$gN_s+1=gN_s+2=\cdots=gN_s+\infty$$

根据上述关系，DMC 自由响应一般表达式为

$$y^*_{t+p}=y_t-\sum_{k=1}^{N_s}(g_k-g_{k+p})\Delta u_{t-k} \tag{6.17}$$

现在自由响应包含当前过程输出 y_t，而这也包括在 DMC 控制算法内，见式（6.12）。

例 6.3.1　考虑下面的一阶过程的 DMC 设计：

$$G(z^{-1})=\frac{0.3z^{-1}}{1-0.4z^{-1}}$$

选择 $N_s=3$，PH，$N=2$，控制动作数 $M=1$，输出加权矩阵 $\mathscr{Q}=\boldsymbol{I}$，控制权重 $\mathscr{R}=\boldsymbol{I}$。

- 得到控制动作 Δu_t。
- 导出离散控制器转换为函数 u_t/ϵ_t，其中 $\epsilon_t=r_t-y_t$，假定设置点为一阶的，即 $r_t=r_{t+1}=r_{t+2}=\cdots$。
- 模拟单位阶跃设定点变化的闭环响应。

已知：$G(z^{-1})=\dfrac{0.3z^{-1}}{1-0.4z^{-1}}$、$N_s=3$、$N=2$、$M=1$、$\mathscr{Q}=\boldsymbol{I}_{2\times 2}$和 $\mathscr{R}=1$，希望由此得到 Δu_t。

解决方案的第一步是要计算阶跃响应系数，它可以由下面的表达式确定：

$$y(z^{-1})=G(z^{-1})\frac{1}{1-z^{-1}}=\frac{0.3z^{-1}}{1-0.4z^{-1}}\,\frac{1}{1-z^{-1}}$$

$$=0.3z^{-1}\left[\frac{1.667}{1-z^{-1}}-\frac{0.667}{1-0.4z^{-1}}\right]$$

将 $y(z^{-1})$ 变成时域响应得到

$$y_n=0.3[1.667(1)^{n-1}-0.667(0.4)^{n-1}]\qquad n\geqslant 1$$

考虑采样时间延迟得到

$$y_n=\begin{cases}0 & n<1\\ 0.5001-0.2001(0.4)^{n-1} & n\geqslant 1\end{cases}$$

阶跃响应系数，现在可以通过代入 n 的整数值而计算出：

$n=0\quad y_n=0$

$n=1\quad y_n=g_1=0.3$

$n=2 \quad y_n = g_2 = 0.4201$

$n=3 \quad y_n = g_3 = 0.4681$

$n=4 \quad y_n = g_4 = 0.4873$

$n=5 \quad y_n = g_5 = 0.4950$

$\vdots$

由阶跃响应系数可将动态矩阵写为

$$\boldsymbol{G} = \begin{bmatrix} 0.3 \\ 0.4201 \end{bmatrix}$$

计算得到控制增益矩阵：

$$\boldsymbol{k}_c = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{G} + \boldsymbol{R})^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}} = [0.2369 \quad 0.3316]$$

由调节时间 $N_{\mathrm{s}} = 3$ 得

$$y^*_{t+p} = y_t - \sum_{i=1}^{3} (a_i - a_{i+p})\Delta u_{t-i} \quad p = 1, 2$$

即

$$\begin{cases} y^*_{t+1} = y_t + 0.1201\Delta u_{t-1} + 0.048\Delta u_{t-2} + 0.0192\Delta u_{t-3} \\ y^*_{t+2} = y_t + 0.1681\Delta u_{t-1} + 0.0672\Delta u_{t-2} + 0.0269\Delta u_{t-3} \end{cases}$$

最后导出控制增量行动：

$$\Delta u_t = \sum_{i=1}^{2} k_{1i}(r_{t+i} - y^*_{t+i}) = k_{11}(r_{t+1} - y^*_{t+1}) + k_{12}(r_{t+2} - y^*_{t+2})$$

代入数值并注明 $r_t = r_{t+1} = r_{t+2}$ 以及 $\epsilon = r_t - y$，得到

$$\begin{aligned} \Delta u_t &= 0.2369(r_t - y_t - 0.1201\Delta u_{t-1} - 0.048\Delta u_{t-2} - 0.0192\Delta u_{t-3}) \\ &\quad + 0.3316(r_t - y_t - 0.1681\Delta u_{t-1} - 0.0672\Delta u_{t-2} - 0.0269\Delta u_{t-3}) \\ &= 0.2369\epsilon_t - 0.0284\Delta u_{t-1} - 0.0114\Delta u_{t-2} - 4.55\times 10^{-3}\Delta u_{t-3} \\ &\quad + 0.3316\epsilon_t - 0.0557\Delta u_{t-1} - 0.0223\Delta u_{t-2} - 8.92\times 10^{-3}\Delta u_{t-3} \end{aligned}$$

使用逆变换 z^{-1}，并重排方程得

$$\begin{aligned} (1 - z^{-1})u_t &= 0.5685\epsilon_t - 0.0841(1 - z^{-1})z^{-1}u_t - 0.0337(1 - z^{-1})z^{-2}u_t \\ &\quad - 0.0135(1 - z^{-1})z^{-3}u_t \end{aligned}$$

控制规则用传递函数形式最终表示为

$$\frac{u_t}{\epsilon_t} = \frac{0.5685}{1 + 0.0841z^{-1} + 0.0337z^{-2} + 0.0135z^{-3}}$$

当在设定点上发生单位阶跃变化时，闭环响应的仿真结果由图 6.2 给出。

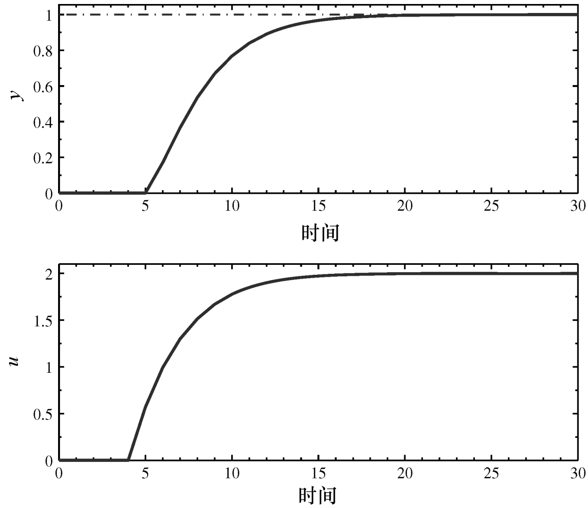


图 6.2　仿真 DMC 闭环响应，上图是输出，下图是控制动作

6.4　NMPC

在过去的几十年中，线性 MPC 由于其易于实施和对工艺过程限制的可控性，一直深受学术界和工业广泛研究。另一方面，由于非线性 MPC（NMPC）计算负荷繁重，实践中相对于线性 MPC 的关注较少。此外，非线性模型的可用性小、非离散化的复杂性高和状态的估计难，这都局限了 NMPC 的应用。但是，廉价的计算能力有新的进展，这逐步削弱了其局限性，并在 NMPC 中的应用中产生了广泛的兴趣。该过程在数学上可叙述如下。

阶跃响应模型或传递函数模型通常用于线性 MPC，NMPC 的自然模型是状态空间模型。其模型表示为

$$x_{t+1} = f(x_t, u_t, w_t) \tag{6.18}$$

$$y_t = g(x_t, v_t) \tag{6.19}$$

式中，x、u 和 w 分别是状态、输入和扰动；状态和输出函数 f 和 g 可以采取线性和非线性的形式。

MPC 的公式化可以写成一组用未来的控制动作 u_t，u_{t+1}，…，u_{t+M-1}来求解以下优化问题：

$$\min_{u_t, u_{t+1}, \cdots, u_{t+M-1}} J = \sum_{i=1}^{N} \left[\| \hat{x}(t+i \mid t) - x_{t+i}^{\text{ref}} \|_{\boldsymbol{Q}}^2 + \| u_{t+i} - u_{t+i}^{\text{ref}} \|_{\boldsymbol{R}}^2 + \| \Delta u_{t+i} \|_{\boldsymbol{S}}^2 \right] \tag{6.20}$$

约束条件为

$$x_{\min} \leqslant \hat{x}(t+i \mid t) \leqslant x_{\max} \tag{6.21}$$

$$u_{\min} \leqslant u_{t+i} \leqslant u_{\max} \tag{6.22}$$

$$\Delta u_{\min} \leqslant \Delta u_{t+i} \leqslant \Delta u_{\max} \tag{6.23}$$

式中，$\hat{x}(t+i \mid t)$ 为预测状态；x_{t+i}^{ref}为状态的参考量；u_{t+i}^{ref}为输入的参考量；N 代表 PH，其中未来状态被测量，目标函数得到最小化；M 为控制域且其控制行动得到最优化。

为了其具有可行性，$M \leqslant N$，控制动作为常数，即 $\Delta u_{t+i}=0$（$i \geqslant M$）。需要注明的是需要考虑控制信号 $\| u_{t+i}-u_{t+i}^{\text{ref}} \|_{\boldsymbol{R}}^2$以及回转率 $\| \Delta u_{t+i} \|_{\boldsymbol{S}}^2$的限制，这是在实际应用中 MPC 包括线性 MPC 的情况。然而为了方便数学推导，如先前讨论过的 DMC 示例中，$\| u_{t+i}-u_{t+i}^{\text{ref}} \|_{\boldsymbol{R}}^2$通常需要下调。现在一旦控制动作已知，只有 u_t 得到应用，在时间 $t+1$ 时重复上述步骤。

在上述 MPC 公式中，目标函数表达了其状态变化和本征二次方程。MPC 公式可以直接表示输出变量并写为

$$\min_{u_t, u_{t+1}, \cdots, u_{t+M-1}} J = \sum_{i=1}^{N} \left[\| g(\hat{x}(t+i \mid t), w_t) - y_{t+i}^{\text{ref}} \|_{\boldsymbol{Q}}^2 + \| u_{t+i}-u_{t+i}^{\text{ref}} \|_{\boldsymbol{R}}^2 + \| \Delta u_{t+i} \|_{\boldsymbol{S}}^2 \right] \tag{6.24}$$

为了实现上述 NMPC 算法，需要解决非凸最优化问题。但由于约束条件的存在，有时这很难实现，特别是对于更复杂的系统来说实现起来更为困难。这个困难是因为解决方法必须与已选择的采样时间都准备好。获得非线性离散模型后，（如 6.6 节将叙述的）标准 NMPC 工作流程可总结如下：

1）从系统中获取当前测量数据。

2）使用非线性离散模型估计系统的不可测状态。

3）计算一组当前和未来的最佳输入，通过 PH 在没有反约束的条件下，使目标函数最小化。

4）实现最佳输入的步骤 1。

5）重复在每个采样间隔步骤 1~4。

在系统的复杂性的基础上，线性 KF、EKF、UKF 或不同变体的粒子过滤器，可以根据步骤 2，估算不可测状态。对于复杂的非线性系统，如 SOFC（固体氧化物燃料电池）系统，EKF 的实现可能变得繁琐，也可能不会达到预期的性能。另一方面，该 UKF 提供了不可测状态的更合适的估计，但是需要做一些调整。如果没有适当的调整，它可能无法收敛。

从理论上讲，人们想在凸优化方案中进行无限地预测和控制时域以实现稳定。然而，这将导致更高的计算负荷和计算量，这可能在下一个采样的实例进入之前来不及实现最优控制输入。出于这个原因，需要选择适当长度的有限预测和控制时域。

6.5　NMPC 通用优化准则

NMPC 的调整是影响控制系统强劲性能的最重要因素之一。在式（6.24）的描述中，NMPC 目标函数有 3 个主要参数：输出权重矩阵 $\boldsymbol{Q}$、输入权重矩阵 $\boldsymbol{R}$ 和在偏差输出对操纵变量中的限制矩阵 $\boldsymbol{S}$。根据这些矩阵的权重，在目标函数中的某些项可能会得到更多的关注。此外，采样时间、离散化方法的选择和状态估计技术都在 NMPC 的表现中起主要作用。适当地调整参数可以确保其稳定性。另一方面，错误的选择可能会导致其不稳定、更高的计算负荷和不可行的解决方案。

有关线性 MPC 调整准则的文献报道有很多，然而由于非线性系统分类广泛，所以对于 MPC 非线性调整没有硬性和约束规则。在线性 MPC 准则的基础上，可以考虑下面的规则调整 NMPC：

- 采样间隔：采样间隔应足够小，以充分捕获系统的动态。作为一般的经验法则，采样周期应在 0.05 ~ 0.1s。对于多变量系统，采样间隔应根据所有过程变量的快速动态计算。
- 预测域：作为一般的经验法则，PH 应该能够反应这个过程的稳态点。PH 小于临界长度可能会导致一个不稳定的闭环系统。而另一方面，较大的 PH 可能会产生更快响应，以及更积极的控制动作和超调量。具有较长预测时域的 MPC 的实现也受到计算能力的限制。对于多变量系统，PH 应基于所有过程变量的慢速动态，对于复杂的非线性系统，诸如对流回路反应器中表现的混乱行为，PH 的选择是特别困难的。对于这种情况，PH 的选择应根据经验和严格的过程模拟来确定。
- 控制时域：一般情况下，相对于 PH 的小控制域导致其对干扰不敏感，产生更少的激烈的响应。3 ~5 个控制时域通常足以满足大多数情况。对于多变量的情况，所述控制域应反映最慢的子系统的动态。
- 加权：控制的变量权重和操纵变量限制应根据变量的大小顺序和客观性来选择。例如，对于同样重要的过程输出变量，权重应选择来衡量过程变量扩展到了类似的幅度。一旦幅值被设置为相同的比例，更多的权重可以根据变量目标的重要性进行应用。

6.6　离散模型：正交配置法

作为第一性原理非线性模型是本质上连续的，应用线性和非线性 MPC 到非线性系统需要额外的步骤。对于线性 MPC，连续模型表达为

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x,u,w) \tag{6.25}$$

在工作点的周围，其可以被转换成离散的线性模型，这样它采用以下形式：

$$x_{t+1} = Ax_t + B_u u_t + B_w w_t \tag{6.26}$$

$$y_t = Cx_t + D_w w_t \tag{6.27}$$

虽然离散线性模型往往足以达到预期的控制效果，但是重要信息往往在线性化的过程中丢失。例如，设计一定容量的蒸馏塔，通常是围绕它的设计运行条件，最大限度提高效益。因此，一个简单的线性模型，是通过工厂测试或线性化第一性原理模型来识别的，这为基于模型控制的成功实施提供了足够的信息。因此，对于这样的系统，第一种方法是优选的，而不是更详细的非线性模型。但是，对于某些非线性系统，情况并非如此。一个混沌系统，如对流回路反应器，在没有任何外界干扰的存在下可以表现出线性或混沌行为。在另一方面，SOFC 表现出电压突降，以及对电流负载的变化需求。这些系统模型进行线性化，周围额定工作条件下不会提供设计控制器的代表模型。因此，这些模型的离散化需要特殊的技术，以保持该过程的非线性属性。

几种数值方法可用来使非线性模型离散化，如正交配置法、Galerkin 法或以平坦度为基础的技术。在本节中，对正交配置，其可用于转换的非线性连续模型进行了讨论。然而，转换后的模型，不采取明确的离散非线性状态空间形式，但仍然可以与 NMPC 架构轻松集成。以下内容提供了正交配置方法的详细描述。

6.6.1　预测时域 1 的正交配置方法

在正交配置方法中，任何功能可以由一个内插多项式与位于一组正交多项式的根节点来近似（Finlayson，1980a），也就是

$$y(x) = \sum_{i=1}^{N+2} b_i P_{i-1}(x) \tag{6.28}$$

式中

$$P_m(x) = \sum_{j=0}^{m} c_j x^j \tag{6.29}$$

第 m 个多项式如下：

$$\int_a^b W(x) P_k(x) P_m(x) \mathrm{d}x = 0 \qquad k = 0,1,2,\cdots,m-1$$

式中，$W(x)$ 是加权函数，见表 6.1。

这里，多项式 P_m 在区间 [a，b] 上时具有 m 个根，因此，如图 6.3 所示，用户不需要随意选择配置点。这优于传统的配置方法，因为对于没有经验的用户，传统的配置方法很容易选择不当节点，会得到错误近似函数。典型地，积分范围取为 0 ~ 1 来归纳问题。式（6.28）和式（6.29）合并后得

$$y(x_j) = \sum_{i=1}^{N+2} d_i x_j^{i-1} \tag{6.30}$$

倒数也可以通过正交多项式来近似，最后，得到了以下形式：

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x_j) = \sum_{i=1}^{N+2} d_i (i-1) x_j^{i-2} \tag{6.31}$$

表 6.1　多项式的根 x_j 和 m 个配置点的加权函数 W_j

m	x_j	W_j
1	0.50000 00000	0.66666 66667
2	$\begin{pmatrix}0.2133248654\\0.7886751346\end{pmatrix}$	$\begin{pmatrix}0.5000000000\\0.5000000000\end{pmatrix}$
3	$\begin{pmatrix}0.1127016654\\0.5000000000\\0.8872983346\end{pmatrix}$	$\begin{pmatrix}0.2777777778\\0.4444444444\\0.2777777778\end{pmatrix}$

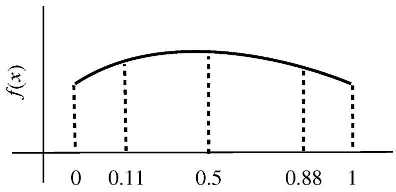


图 6.3　在超前预测中通过由 3 个配置点得到的近似函数

$$\frac{d^2y}{dx^2}(x_j) = \sum_{i=1}^{N+2} d_i(i-1)(i-2)x_j^{i-3} \tag{6.32}$$

用矩阵表示法得

$$y = \boldsymbol{Qd} \qquad \frac{dy}{dx} = \boldsymbol{Cd} \qquad \frac{d^2y}{dx^2} = \boldsymbol{Dd}$$

式中

$$Q_{ji} = x_j^{i-1}$$
$$C_{ji} = (i-1)x_j^{i-2}$$
$$D_{ji} = (i-1)(i-2)x_j^{i-3} \tag{6.33}$$

因此

$$\frac{dy}{dx} = \boldsymbol{CQ}^{-1}y \equiv \boldsymbol{A}_y \tag{6.34}$$

$$\frac{d^2y}{dx^2} = \boldsymbol{DQ}^{-1}y \equiv \boldsymbol{B}_y \tag{6.35}$$

利用具有 3 个配置点的多项式导数的趋近式，其一阶连续模型表示为

$$\frac{dx}{dt} = f(x,u) \tag{6.36}$$

可以通过下面的非线性代数方程来表示：

$$\boldsymbol{A}_0\boldsymbol{X} = \boldsymbol{F}(\boldsymbol{X},u) \tag{6.37}$$

式中

$$\boldsymbol{X}=\begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ x_{3,1} & x_{3,2} & \cdots & x_{3,n} \\ x_{4,1} & x_{4,2} & \cdots & x_{4,n} \\ x_{5,1} & x_{5,2} & \cdots & x_{5,n} \end{pmatrix} \qquad \boldsymbol{F}(\boldsymbol{X},u)=\begin{pmatrix} x_{\text{init}}^{\mathrm{T}} \\ f^{\mathrm{T}}(x_{2,*}^{\mathrm{T}},u,t_2) \\ f^{\mathrm{T}}(x_{3,*}^{\mathrm{T}},u,t_3) \\ f^{\mathrm{T}}(x_{4,*}^{\mathrm{T}},u,t_4) \\ f^{\mathrm{T}}(x_{5,*}^{\mathrm{T}},u,t_5) \end{pmatrix} \tag{6.38}$$

式中，$x_{i,j}$的第一个下标表示该配置点的采样时间间隔，第二下标指状态；x_{init}是状态的初始条件；u 在时间间隔内是恒定的；t 是时间变模型需要的时间变量，但是，对于时不变模型可以省略。

配置点和 $\boldsymbol{A}$ 矩阵列于表 6.2。矩阵 $\boldsymbol{A}_0$ 可以考虑初始条件情况下，通过已知的节点 x_j 的位置和通过（1 0 0 0 0）替换 $\boldsymbol{A}$ 的第一行来计算。

表 6.2 给出了不同配置点［0，1］区间的矩阵。考虑到采样间隔不止是 1，下列更改到动态约束转换成代数约束：

$$\boldsymbol{A}_0\boldsymbol{X}=\Delta t'\boldsymbol{F}(\boldsymbol{X},u) \tag{6.39}$$

表 6.2　从式（6.33）得出的 *m* 点正交配置矩阵

m	$\boldsymbol{A}$
1	$\begin{pmatrix} -3 & 4 & -1 \\ -1 & 0 & 1 \\ 1 & -4 & 3 \end{pmatrix}$
2	$\begin{pmatrix} -7 & 8.196 & -2.196 & 1 \\ -2.732 & 1.732 & 1.732 & -0.7321 \\ 0.7321 & -1.732 & -1.732 & 2.732 \\ -1 & 2.196 & -8.196 & 7 \end{pmatrix}$
3	$\begin{pmatrix} -13 & 14.79 & -2.67 & 1.88 & -1 \\ -5.32 & 3.87 & 2.07 & -1.29 & 0.68 \\ 1.5 & -3.23 & 0 & 3.23 & -1.5 \\ -0.68 & 1.29 & -2.07 & -3.87 & 5.32 \\ 1 & -1.88 & 2.67 & -14.79 & 13 \end{pmatrix}$

6.6.2　预测时域 *N* 的正交配置方法

由式（6.39）中给出的公式可以用于将非线性连续时间动态方程转变为非线性代数方程，从而其可作为 PH 为 1 的等式约束条件用于 MPC。对于 PH 为 N，正交配置的近似扩展可以通过采用下式实现（见图 6.4）：

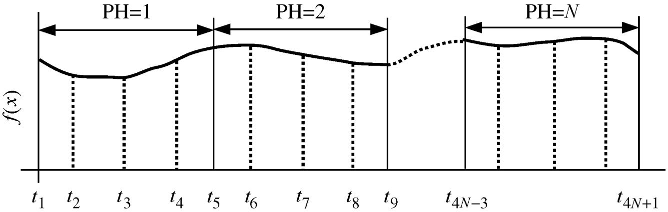


图 6.4　采用三点式配置法和预测时域的函数多项式趋近结果，其中 PH = **N**

$$\tilde{\boldsymbol{A}}\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{X}},\boldsymbol{U}) \tag{6.40}$$

式中

$$\tilde{\boldsymbol{X}} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{4N,1} & x_{4N,2} & \cdots & x_{4N,n} \\ x_{4N+1,1} & x_{4N+1,2} & \cdots & x_{4N+1,n} \end{pmatrix}$$

$$\tilde{\boldsymbol{F}}(\tilde{\boldsymbol{X}},\boldsymbol{U}) = \begin{pmatrix} x_{\text{init}}^{\mathrm{T}} \\ f^{\mathrm{T}}(x_{2,*}^{\mathrm{T}},u,t_2) \\ \vdots \\ f^{\mathrm{T}}(x_{4N,*}^{\mathrm{T}},u,t_{4N}) \\ f^{\mathrm{T}}(x_{4N+1,*}^{\mathrm{T}},u,t_{4N+1}) \end{pmatrix}$$

式中，$X_{i,j}$的第一个下标表示配置点，第二下标指状态；u 沿各 PH 不变；$\tilde{\boldsymbol{A}}$ 由 $\boldsymbol{A}$ 和 $\boldsymbol{A}_0$ 被定义如下：

$$\tilde{\boldsymbol{A}} = \begin{pmatrix} \boxed{\boldsymbol{A}_0} & & & & & \\ & \boxed{\boldsymbol{A}} & & & & \\ & & \boxed{\boldsymbol{A}} & & & \\ & & & \ddots & & \\ & & & & \boxed{\boldsymbol{A}} & \\ & & & & & \boxed{\boldsymbol{A}} \end{pmatrix} \tag{6.41}$$

现在，该模型已被从连续时间的非线性常微分方程（ODE）转换为一组离散的代数方程组，它可以作为等式约束用于 MPC 公式中：

$$\min_{u_t,u_{t+1},\cdots,u_{t+M-1}} J = \sum_{i=1}^{N}\left[\,\|\hat{x}(t+i\mid t)-x_{t+i}^{\text{ref}}\|_{\boldsymbol{Q}}^{2}+\|u_{t+i}-u_{t+i}^{\text{ref}}\|_{\boldsymbol{R}}^{2}+\|\Delta u_{t+i}\|_{\boldsymbol{S}}^{2}\right] \tag{6.42}$$

这归因于

$$\tilde{A}\tilde{X}=\tilde{F}(\tilde{X},U) \tag{6.43}$$

$$x_{min}\leqslant x(t+i\mid t)\leqslant x_{max} \tag{6.44}$$

$$u_{min}\leqslant u_{t+i}\leqslant u_{max} \tag{6.45}$$

$$\Delta u_{min}\leqslant \Delta u_{t+i}\leqslant \Delta u_{max} \tag{6.46}$$

正交配置方法的主要优点是，它不需要选择节点，从而不容易出现人为错误。另外，如上所述，离散模型可以直接在非线性优化 MPC 方案中使用。另一方面，若干内部配置点的选择，对真系统的近似产生深远的影响，该系统再次变化从系统到系统的近似。每个采样间隔之间的节点的数目增加，状态的数量也增加。这反过来又增加了 NMPC 计算量。从混乱系统，连续搅拌釜式反应器到 SOFC 系统，对那些非线性系统进行了广泛仿真模拟，可以发现，在大多数情况下，3 个内部配置点可能就足够接近系统行为。

6.7 MPC 的优、缺点

尽管 NMPC 比线性 MPC 具有优越的性能，NMPC 的实施却不是简单的。首先，它通常需要一个非线性第一性原理模型，其开发和维护需要很多时间和精力。此外，在设备中的任何显著改变也需要校正模型，这可能会需要持续的努力。另一方面，从长远来看线性模型，可以通过执行简单的设备测试来进行更新。

此外，根据不同的模型的复杂性、采样时间、使用非线性状态估计器和控制器的执行周期可以来估计其不可测状态。NMPC 的计算时间比线性 MPC 要长得多。因为实现平台必须能够预先在指定的期限内解决约束优化问题，所以这对 NMPC 的实现造成了困难。这个执行时间依赖于系统的动态特性，并减慢了系统速度。如果控制动作的计算速度不够快，效果就会很差。但随着廉价计算能力的出现，这个问题应该会得到缓解。

总体而言，无论线性还是非线性，MPC 的应用都会有利弊。以下是一些 MPC 的主要优点和缺点：

优点：

● 它可以轻松地处理多变量问题。

● 它可以处理限制，并允许更多的紧密合作过程中的制约因素。其实，这是 MPC 或 NMPC 如何实现其大部分收益的问题。

● 算法的反馈性质可以利用该模型精度中的不确定性来达到某种程度。

● 无需解耦或延迟补偿。它可以用于较慢的系统。然而，对于快速系统，NMPC 受到该平台的计算限制而无法进行。

● 自由度的有效利用。它允许设计非方形系统（输出多于输入），并且根据所分配的优先级和权重，只有一些变量可由 MPC 控制。

缺点：

● 对于较大的系统中，计算负荷可能会很重，这是特别适用于 NMPC 的情况。

● 在结构的复杂性和调优参数的数量的基础上，调整是很困难的。

● 它需要一个明确的模型。控制器取决于如何准确地确定模型，这造成了一定性能的限制。而且，随着这个过程中的结构变化，该模型可能会变得无效，并且需要被重新识别。

6.8　最优化

无论在汽车行业、石油天然气部门还是污水处理厂，业务受经济驱动。企业的主要目标是提高产量，在不断增长的全球竞争中降低运营成本生存下来也是很重要的。设计合理的优化器可以通过降低成本、实现利润最大化来实现经济目标。在不违反约束条件下，它可以提供最佳的最优解决方案，帮助实现经济目标。由于工艺设备的可靠性和完整性的限制，真正的最优经济的解决方案往往超出约束的限制。因此，对于大多数的情况，优化器的工作原理类似于一个约束推进器，并帮助找到定位在哪里，这样可以使大部分的利润（或成本）最小化。

线性 MPC 或 NMPC 还可以作为优化方案，其在原有的控制目标中为经济损失函数 $L(x, u)$，如式（6.20）所示，即

$$\begin{aligned}\min_{u_t,u_{t+1},\cdots,u_{t+M-1}} J = \sum_{i=1}^{N} [\, \| \hat{x}(t+i\,|\,t) - x_{t+i}^{\mathrm{ref}} \|_{\boldsymbol{Q}}^2 + \| u_{t+i} - u_{t+i}^{\mathrm{ref}} \|_{\boldsymbol{R}}^2 + \| \Delta u_{t+i} \|_{\boldsymbol{S}}^2 \,] \\ + \boldsymbol{W} \sum_{i} L(\hat{x}(t+i\,|\,t), u_{t+i})\end{aligned} \tag{6.47}$$

式中，$L(\hat{x}(t+i\,|\,t), u_{t+i})$ 是经济损失；$\boldsymbol{W}$ 是该损失的权重；$\boldsymbol{Q}$、$\boldsymbol{R}$、$\boldsymbol{S}$ 和 $\boldsymbol{W}$ 为目标函数中的加权矩阵，将决定系统是否是动态控制或进行了优化的稳定状态。

这个构想有几个问题：首先，在每个实例中，都存在经济优化，即使系统大多数情况在稳定状态下工作；第二，从控制和优化得到的方案可产生冲突的目标。

另一方面，如果该方案在分层结构处于分离状态，这两个目标在其中任何一方的情况下，不可能实现。那么 NMPC 可以通过优化器来驱动该工厂计算出最优的运行状态［即提供 x^{ref} 给式（6.20）］。总之，高级别优化如下所示：

$$\min_{x_{\mathrm{ref}}} J = \sum_{i=1}^{N} L(x, u) \tag{6.48}$$

$$f_s(x, u, w) = 0 \tag{6.49}$$

$$x_{\min} \leqslant \hat{x} \leqslant x_{\max} \tag{6.50}$$

$$u_{\min} \leqslant u \leqslant u_{\max} \tag{6.51}$$

$$\Delta u_{\min} \leqslant \Delta u \leqslant \Delta u_{\max} \tag{6.52}$$

式中，f_s（x，u，w）是该系统的非线性稳态模型。

优化器可以在很慢的频率下运行，或根据需要运行。NMPC 可以在更高频率下运行，它可以在不违反约束条件下动态地推动工厂向稳态优化目标运行。总体方案由图 6.5 来描述。在第 12 章将给出分层优化的例子。

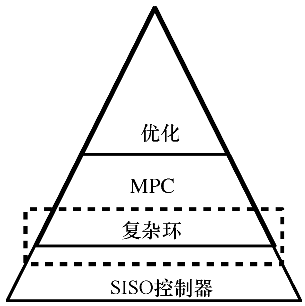


图 6.5　控制层次结构

6.9　范例：混沌系统

在物理系统中混沌是常见的，减少混乱可以使系统性能得以改善。化学加工工业热的环流反应器对流，可以通过一个相对简单的非线性动态模型来表示。根据不同的加热速率，它可以显示稳定的、永久的振荡或混沌行为。该反应器描述的常微分方程为

$$\dot{x}_1 = \alpha(-x_1 + x_2) \tag{6.53}$$

$$\dot{x}_2 = -x_2 - x_1 x_3 \tag{6.54}$$

$$\dot{x}_3 = x_1 x_2 - x_3 - \beta \tag{6.55}$$

式中，α 类似于普朗克数；β 为升温速率，它直接与温度差 ΔT 成正比，并等价于瑞利数；状态 x_1，x_2 和 x_3 正比于平均截面循环内速度、沿水平方向和沿垂直方向的温度差。

基于升温速率（β），对流回路反应器可以显示稳定或混乱的行为。

对流环管反应器的复杂的动态性质（混沌行为）有利于测试 NMPC。应用在对流循环 NMPC 呈现混沌行为，其中 UKF 作为状态估计。对于该仿真，已经假设对流循环仅在水平温差来测量，其他两种状态利用 UKF 估计（垂直温差和管内流量）。测量噪声也被添加到水平温差中。来自 UKF 估计的输出被用于计算 NMPC 的未来控制动作。必须指出，设备是利用第一性原理连续模型在 MATLAB® 下仿真的，而正交配置方法被用在 NMPC 公式中来呈现离散模型。因此，控制方案中提供了设备模型不匹配情况。其结果示于图 6.6 中。从仿真结果看，NMPC 可以使混沌系统稳定。时间的稳定在很大程度上取决于限制功能、采样率和不同模仿真实生

活的场景的约束条件。

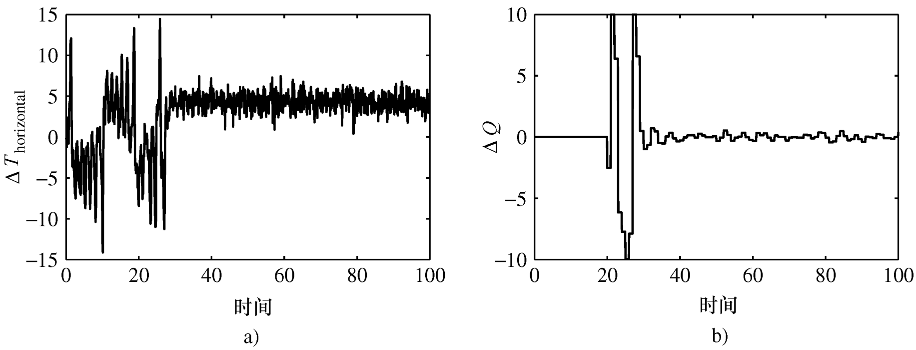


图 6.6　由 MPC 连同 UKF 作为状态估计量实现对流环流反应器的控制
a）横向温差（测量输出）　b）升温速率

6.10　注解和参考文献

近些年，MPC 成为非常成功的控制技术。Garcia 等人（1989）、Morari 和 Lee（1999）对 MPC 都有详细的研究评论。2000 年，Rawlings 提供了理论和 MPC 应用的一个不错的入门教程。研究发现，Allgower 等人（2004）、Kouvaritakis 等人（2000）、Rao 和 Rawlings（2000）分别对非线性模型预测控制和滚动时域状态估计进行了更深一步的讨论。

众所周知，线性 MPC 也被广泛地引用。动态矩阵控制的 DMC 的理论部分都是以 Huang 和 Kadali（2008）（施普林格出版社）为基础的。线性 MPC 调整准则在相关文献上非常常见，如 Lee 和 Yu（1994）和 Oshima 等人（1991）所述。Shridhar 和 Cooper（1998）也提供了一个非常好的 DMC 调整准则，它可以很容易地应用到线性 MPC 或 NMPC。

对于一个非线性连续时间模型，一个重要步骤是近似一个离散的模型从而达到最佳服务的目的。6.6 节提供了关于正交配置方法的简要说明。详细的结论参见 Finlayson（1980b）、Henson 和 Seborg（1997）、Murshed 等人（2003）、Villadsen 和 Michelsen（1978）的研究。

由于处于范围之外，优化技术没有在书中详细探讨。相反，NMPC 优化的整合已经从控件层次结构的角度开始讨论。Rao（2009）为探索不同的优化技术提供了一个良好的开端。

第 2 部分　管式 SOFC

第 7 章　管式 SOFC 动态模型：第一性原理方法

7.1　SOFC 堆栈的设计

一个典型的单一燃料电池只能产生大约 0.7V 的电动势，所以许多电池必须堆放在一起得到一个更高的输出电压。那么设计一个 SOFC 堆栈就需要解决叠加引起的问题，如电导率、气密封和热应力。因此，一个 SOFC 堆栈的配置已成为主要的研究和开发工作。

SOFC 堆栈配置可以分为两类：平板式和管式。材料的强度和耐久性限制了 SOFC 组件的力学性能。在兆瓦级的应用中，管式 SOFC 堆栈更有前景，因为相对于一个大的平板来说，热能与机械应力对管的强度的影响更小。因此，管式设计大大降低了所需要的 SOFC 电解质材料的强度要求。它也可以避免使用气体密封，从而保持阳极和阴极燃料通道分开。另一方面，平面设计中必须采用气体密封，而气体密封在高温下是一个挑战。

堆栈设计如图 7.1 所示，是一个典型的管式设计。这个设计包含大量相似的电池管组合。每个组合包括两个管：电池管和注入管。外管是 SOFC 单元。它是一个空气电极支撑管，电解液沉积在上面，紧随其后的是阳极。内管是空气注入和引导管，由氧化铝制成，预热空气可以由此注入电池管的底部，然后从注入管和电池管的间隙中流过阴极表面。燃气通过电池之间的空隙流到阳极表面。

在本章，以一个单独的电池管为模型进行详解。主要的目标是通过模型了解 SOFC 的过程，同时，阐明了第 2 章中所介绍的化学工程原理是如何应用于开发 SOFC 详细的动态模型中的。以电化学转换过程开始，然后是扩散过程，扩散过程把反应物由反应体传输到反应位点。然后考虑反应体在流动通道中的传输过程。另外，对详细的动态现象，如副反应和温度的影响，进行了研究。SOFC 过程所涉及的物理过程的第一性原理模型将被转换为一个状态空间模型。最后，这个模型通过模拟用于研究动态性能中。

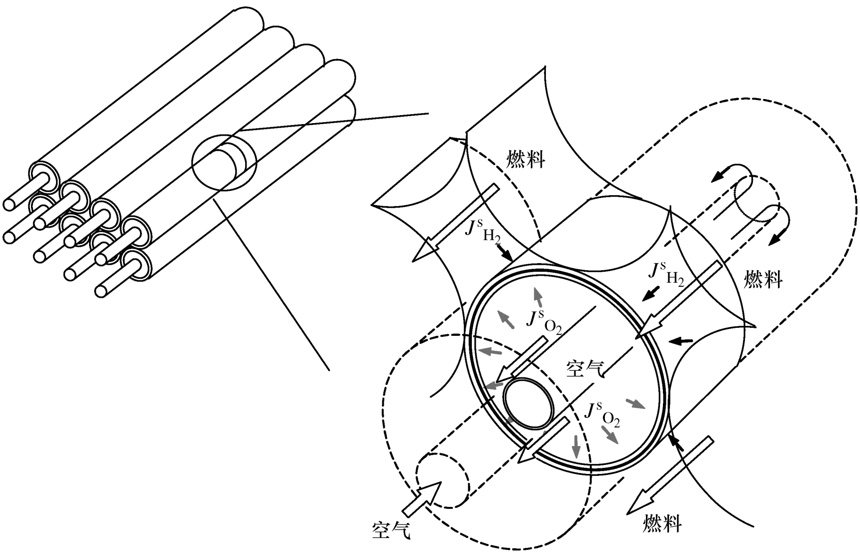


图 7.1　管式 SOFC 的例子

7.2　转换过程

在做 SOFC 模型的过程中，首先要考虑的是转换过程。这个转换过程指的是外部负载的电化学反应，它建立了电池阴极和阳极的电势差，供应持续的电子流，称为电流。因此，转换过程是 SOFC 过程的核心步骤。图 7.2 在微观范围内说明了转换过程的原理。

7.2.1　电化学反应

在电池产生电流的过程中，直接涉及的电化学反应如下：

$$\text{阳极}\quad H_2 + O^{2-} \rightarrow H_2O + 2\,e^-$$

$$\text{阴极}\quad \frac{1}{2}O_2 + 2\,e^- \rightarrow O^{2-} \tag{7.1}$$

$$\text{阳极}\quad CO + O^{2-} \rightarrow CO_2 + 2\,e^-$$

$$\text{阴极}\quad \frac{1}{2}O_2 + 2\,e^- \rightarrow O^{2-} \tag{7.2}$$

大量的研究表明，阳极反应是很快的。在合适的催化剂下，阴极反应是相对快的，并不是一个限制步骤。事实表明，这些电化学反应是相当快的，足以把到达反应位点的所有反应物反应掉。在这些观察结果的基础上，可以做这样一个假设，产生电动势（EMF）和电流的这些电化学反应是瞬间的。这样就可以忽略掉电化学反应动态，只考虑稳态。

电化学反应的结果，即 EMF，是受几种因素影响的，主要因素是反应物的分

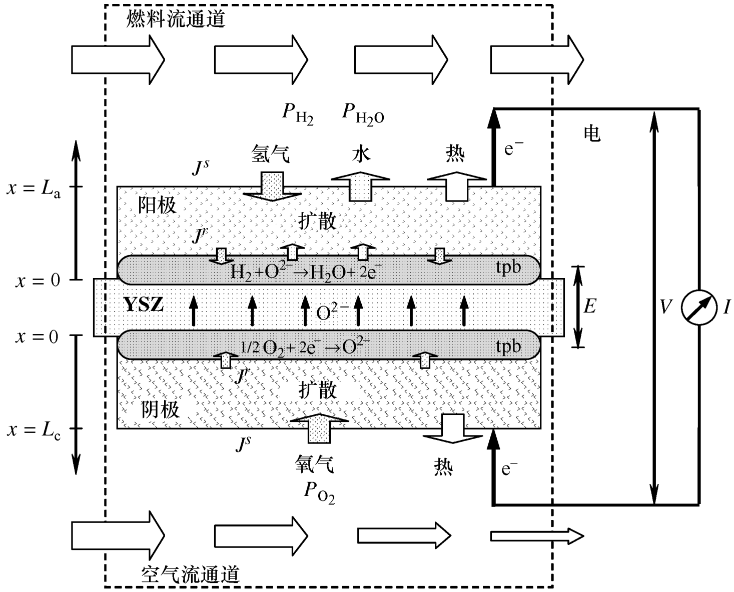


图 7.2　SOFC 原理

压，其关系可以用 Nernst 方程来表示（Larminie 和 Dicks，2003a）：

$$E = E^\circ + \frac{RT}{2F}\ln\left(\frac{a_{H_2}\cdot a_{O_2}^{0.5}}{a_{H_2O}}\right) \tag{7.3}$$

式中，a_{H_2}、a_{O_2}和 a_{H_2O}是相应物质的活性。在多数低压状况下，可以合理地假设气体是理想气体，那么活性可以表示为

$$a_{H_2} = \frac{p_{H_2}}{p^\circ}$$

$$a_{O_2} = \frac{p_{O_2}}{p^\circ}$$

$$a_{H_2O} = \frac{p_{H_2O}}{p^\circ_{H_2O}}$$

式中，p_{H_2}、p_{O_2}和 p_{H_2O}是反应物分压；P°是标准压力；$P^\circ{}_{H_2O}$是某温度下的蒸汽压力。

E 也称为 Nernst 电压。如果 $P^\circ = P^\circ{}_{H_2O} = 1\text{atm}$，那么 Nernst 方程可以化简为

$$E = E^\circ + \frac{RT}{2F}\ln\left(\frac{p_{H_2}p_{O_2}^{1/2}}{p_{H_2O}}\right) \tag{7.4}$$

电化学反应出现在三相反应界面，在三相反应界面附近影响 EMF 的是分压。Nernst 方程更合适的表达方式为

$$E = E^\circ + \frac{RT}{2F}\ln\left(\frac{p_{H_2}^{tpb}p_{O_2}^{tpb\,1/2}}{p_{H_2O}^{tpb}}\right) \tag{7.5}$$

式中，$p_{H_2}^{tpb}$、$p_{O_2}^{tpb}$和 $p_{H_2O}^{tpb}$是三相反应界面附近的分压。

在电化学反应过程中，反应产生的一部分电势能用来克服活化能来维持反应，这就降低了电化学反应所产生的电势。这部分势能的下降称为活化能损失。活化能损失和输出电流的关系可以用 Butler-Volmer 方程来表示：

$$I = I_0\left\{\exp\left(\beta\frac{nF\eta_{act}}{RT}\right) - \exp\left[-(1-\beta)\frac{nF\eta_{act}}{RT}\right]\right\} \tag{7.6}$$

式中，β 是传导系数；I_0 是交流电流。

在燃料电池中，传导系数通常为 0.5（Chan 等人，2001）。

当 $\beta = 0.5$ 时，活化能损失可以由式（7.6）算出

$$\begin{aligned} \eta_{act,a} &= \frac{2RT}{nF}\sinh^{-1}\left(\frac{I}{2I_{0,a}}\right) \\ \eta_{act,c} &= \frac{2RT}{nF}\sinh^{-1}\left(\frac{I}{2I_{0,c}}\right) \end{aligned} \tag{7.7}$$

交流电流指的是不流经外部电路，但是为了保持内部最低反应，而通过内部电极的电流。这取决于活化能、温度、局部分压，可以由下式计算：

$$\begin{aligned} I_{0,a} &= 7\times10^9 A(p_{H_2}^{tpb}p_{H_2O}^{tpb})\exp\left(-\frac{E_{act,a}}{RT}\right) \\ I_{0,c} &= 7\times10^9 A(p_{O_2}^{tpb})^{0.25}\exp\left(-\frac{E_{act,c}}{RT}\right) \end{aligned} \tag{7.8}$$

式中，$E_{act,a}$和 $E_{act,c}$分别是阳极和阴极的活化能。

如果考虑活化损失，不可逆电压可以定义如下：

$$E = E^\circ + \frac{RT}{2F}\ln\left(\frac{p_{H_2}^{tpb}p_{O_2}^{tpb1/2}}{p_{H_2O}^{tpb}}\right) - \eta_{act,a} - \eta_{act,c} \tag{7.9}$$

式中，E°是源于 Gibb's 自由能的电动势，Gibb's 自由能就是在理想条件下，电化学反应所释放的能量。

通常，E°用一个有关 T（Campanari 和 Iora，2004）的一次多项式表示如下：

$$E^\circ = 1.273 - 2.7645\times10^{-4}T \tag{7.10}$$

因为电化学速率比较快，所以产生电流的速度由相对较慢的反应物供应率来决定。根据第 2 章所介绍的质量平衡原理，通过考虑电子和反应物之间的物料平衡，可以把电流和反应物提供利率之间的关系确定为

$$I = 2FJ_{H_2}^r = 2FJ_{H_2O}^r = 4FJ_{O_2}^r \tag{7.11}$$

式中，上角 r 代表在三相反应界面燃料消耗或水蒸气产量。

电流主要由外部负载所决定，在正常操作条件下，一个外部负载需要或者吸收多少电流，反过来，燃料电池也会通过电化学反应产生同样的电流。结果，反应物消耗速率会自动调节，这反过来就会影响传导过程以及反应位点的局部分压。因此，电流是一个连接阳极反应和阴极反应的协调变量，这样就形成了一个连续的燃

料电池反应过程。它也决定了有多少反应会发生，这些过程的动态模型将会在7.2.2节中讲到。

7.2.2　电流动态

产生的电流需要从三相反应界面传输到内部负载。双层电容和燃料电池的内阻影响着供应到外部负载的电压和电流。

7.2.2.1　内阻

燃料电池的电阻由三部分组成：①电极，阴极和阳极，它们都是纯欧姆电阻；②电解液的离子电阻，可以看作电子电阻；③从电解液到电极的电阻，其中包括电荷转移和晶界电阻。虽然这些电阻的机制不同，但是它们对电压和电流的最终影响是类似的，并且都可以视为欧姆电阻。因为阳极和电解液之间、阴极和电解液之间双层电容的存在，这些电阻的综合效应可以更准确地称作阻抗。这个术语不仅表示了对电路中电流流动的抵抗，也表示了对交流电起阻碍作用。当电流流经固有阻抗时，电压只会减小并不会延迟。

如前所述，SOFC的内阻包括三个电阻和两个电荷电容。因为它们对电压和电流的影响类似于一个*RC*电路，所以固有阻抗经常以等效*RC*电路为模型。等效*RC*电路如图7.3所示。在这里，R_e是电解液的离子电阻，R_a和R_c分别代表阴极和阳极的欧姆和晶界电阻，R_{cta}和R_{ctc}代表了电荷转移电阻，C_{dla}和C_{dlc}代表了阴极、阳极和电解液之间的电荷双层电容。

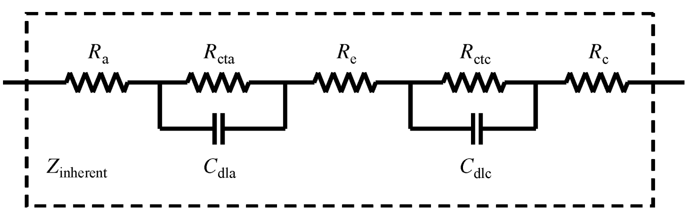


图7.3　内阻的等效电路

电化学阻抗谱（EIS）方法是研究SOFC固有阻抗的一个非常有效的方法，EIS广泛应用于研究动态电化学，包括电力学。从本质上讲，它是一个频域识别法。在不同频率下，用小的交流电刺激燃料电池，看到输出电压，就可以识别内部电化学和电力学参数。

7.2.2.2　等效电路

SOFC内阻的组件如图7.3所示。从理论上讲，阻抗谱图应该显示两个半圆图形（Macdonald，1987）。忽略反应物扩散过程中引起的阻抗效应，因为阳极和阴极反应动态相似，这两个半圆图形可以合并成一个“歪曲”的半圆（Wagner等人，1998）。因此，整个固有阻抗可以用一个*RC*单元来模拟。

电压是在两个电极层之间产生的，在三相反应界面，电极的欧姆电阻会降低电压（见图7.2）。双电容层可以平滑EMF电势改变引起的输出电压的波动。图7.4

所示是一个更合适的 SOFC 等效电路，其中，R_o 是固有阻抗的总欧姆电阻，R_{ct} 是电荷转移电阻，C_{ct} 是电荷转移电容。这 3 个参数可以从一个阻抗谱图来确定。

这个等效电路不同于 Larminie 和 Dicks（2003a）所说的那个，理想电池与一个 *RC* 对串联。那个等效电路对电池电势的变化显示出衍生效应。这意味着，如果一个理想电池的电势跳跃到一个高的水平，输出电压将会有一个激增。这种响应还没有被发现和报道，对于它在燃料电池中的出现，似乎没有清晰的物理解释。因此在这里使用的模型是一种改进的等效电路（见图 7.4）。它平滑了输出电压，同时可以模拟电池内部和外部的运行情况。通过这个简单的等效电路，使用电路原理可以使问题简化。

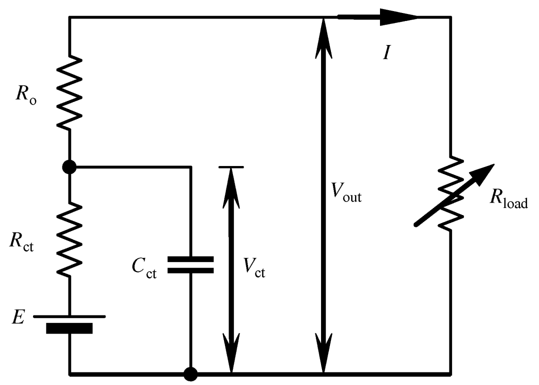


图 7.4　单燃料电池的等效电路

外部负载是决定燃料电池可以提供多大电流的主要因素。因为双电容层的存在，电流也会受到输出电压动态的影响。在这一点上，暂时可以假设反应物的供应率不是限制因素。那么，电压和电流的关系就可以通过基尔霍夫定律来表征：

$$
\begin{aligned}
\dot{V}_{ct} &= \frac{1}{R_{ct}C_{ct}}E - \frac{1}{R_{ct}C_{ct}}V_{ct} - \frac{1}{C_{ct}}I \\
V_{out} &= V_{ct} - IR_o \\
I &= \frac{V_{ct}}{R_o + R_{load}}
\end{aligned}
\tag{7.12}
$$

7.3　扩散动态

现在，来考虑反应物提供过程以及对传输过程的限制。尽管实际的反应过程发生在三相反应界面，但是反应物必须从供应流体，通过多孔电极，到反应位点。因为多孔电极的膜很薄，所以在研究 SOFC 的静态性能时，大多数研究者都忽略了扩散的影响。然而 EIS 研究证明在 SOFC 的动态响应中，扩散过程起着重要的作用。

为了保证输出电流的连续性，从反应物源到反应位点的反应物的连续供应是一个必要条件。这是一个相对简单的传导过程。传导过程包括两个阶段：首先，从流体源到电池表层；然后，通过多孔电极到反应位点（Campanari 和 Iora，2004）。在第一阶段，物质流必须经过界面层到达电池表层。在第二阶段，物质流通过多孔电

极扩散到反应位点。在这两个阶段，扩散都是主要的传输方式。

扩散的驱动力是浓度梯度，见第 2 章说明。描述扩散，最广泛使用的是菲克定律（Fick's law）：

$$j = -D\frac{\mathrm{d}C}{\mathrm{d}x} \tag{7.13}$$

由于人们也很关注扩散动态，所以必须把时间考虑进去，如第 2 章所讲的，通过把菲克定律应用到连续方程中，动态扩散方程可以写为

$$\frac{\partial C}{\partial t} = -D\frac{\partial^2 C}{\partial x^2} \tag{7.14}$$

式中，C 是质量浓度；D 是扩散系数；x 是扩散深度的坐标，如图 7.2 所示。

菲克定律表明，浓度与扩散路径中的位置 x 有关。为了获得三相反应界面的浓度，多数人采用特定元素的方法来测定路径上的浓度分布，然后计算出终端边界的浓度。这里很关注三相反应界面和物质流这两个边界层的关系。在下面的部分，可以用一个更简单的方法解决这个问题。

7.3.1 扩散传递函数

如式（7.14）所示，扩散的动态模型不仅随时间变化，还随空间变化，这就很难用于控制应用领域。在这样的特殊的扩散研究过程中，不关注在整个扩散路径中浓度的动态分布，关注的是在两个边界之间浓度的动态关系。拉普拉斯变换可以解决这个问题。通过拉普拉斯变换，这种动态关系可以用不含空间变量的函数来表示。

对式（7.14）做拉普拉斯变换，可以将偏微分方程变换为关于位置 x 的二阶常微分方程：

$$\frac{\mathrm{d}^2 C(s)}{\mathrm{d}x^2} + \frac{s}{D}C(s) = 0 \tag{7.15}$$

临界条件是

$$\begin{gathered} j^r(s) = -D\left.\frac{\mathrm{d}C(s)}{\mathrm{d}x}\right|_{x=0} \\ C^b(s) = C(s)\big|_{x=l} \end{gathered} \tag{7.16}$$

解式（7.15）得

$$\begin{aligned} C(s)(x) = {} & \frac{C^b(s) + \dfrac{j^r(s)}{\sqrt{Ds}}\exp\left(-\sqrt{\dfrac{s}{D}}l\right)}{\exp\left(\sqrt{\dfrac{s}{D}}l\right) + \exp\left(-\sqrt{\dfrac{s}{D}}l\right)}\exp\left(\sqrt{\dfrac{s}{D}}x\right) \\ & + \frac{C^b(s) - \dfrac{j^r(s)}{\sqrt{Ds}}\exp\left(\sqrt{\dfrac{s}{D}}l\right)}{\exp\left(\sqrt{\dfrac{s}{D}}l\right) + \exp\left(-\sqrt{\dfrac{s}{D}}l\right)}\exp\left(-\sqrt{\dfrac{s}{D}}x\right) \end{aligned} \tag{7.17}$$

假设气体是理想气体，流动面积是 A，在 $x=0$ 的三相反应界面，整理方程后，可以得到三相反应界面附近的局部分压：

$$p^{\text{tpb}}(s) = -\frac{\exp\left(\sqrt{\frac{s}{D}}l\right) - \exp\left(-\sqrt{\frac{s}{D}}l\right)}{\exp\left(\sqrt{\frac{s}{D}}l\right) + \exp\left(-\sqrt{\frac{s}{D}}l\right)} \frac{1}{\sqrt{Ds}} \frac{RT}{A} J^r(s) + \frac{2}{\exp\left(\sqrt{\frac{s}{D}}l\right) + \exp\left(-\sqrt{\frac{s}{D}}l\right)} p^b(s) \tag{7.18}$$

扩散层的表面，流量是$J^s = AD\left.\frac{\mathrm{d}C}{\mathrm{d}x}\right|_{x=l}$。利用浓度 C，可以得到下面的方程：

$$J^s(s) = \frac{2}{\exp\left(\sqrt{\frac{s}{D}}l\right) + \exp\left(-\sqrt{\frac{s}{D}}l\right)} J^r(s) + \frac{\exp\left(\sqrt{\frac{s}{D}}l\right) - \exp\left(-\sqrt{\frac{s}{D}}l\right)}{\exp\left(\sqrt{\frac{s}{D}}l\right) + \exp\left(-\sqrt{\frac{s}{D}}l\right)} \sqrt{Ds} \frac{A}{RT} p^b(s) \tag{7.19}$$

7.3.2　简化的扩散传递函数

在实际应用中，前面部分所讲的函数过于复杂，可以通过省略高阶项进行进一步简化。

$\exp\left(\sqrt{\frac{s}{D}}l\right)$和 $\exp\left(-\sqrt{\frac{s}{D}}l\right)$在零点关于$\sqrt{\frac{s}{D}}l$ 和 $-\sqrt{\frac{s}{D}}l$ 的泰勒级数展开式如下：

$$\exp\left(\sqrt{\frac{s}{D}}l\right) = 1 + \sqrt{\frac{s}{D}}l + \frac{1}{2}\frac{s}{D}l^2 + \frac{1}{6}\frac{s^{\frac{3}{2}}}{D^{\frac{3}{2}}}l^3 + \frac{1}{24}\frac{s^2}{D^2}l^4 + O\left(\sqrt{\frac{s}{D}}l\right)$$

$$\exp\left(-\sqrt{\frac{s}{D}}l\right) = 1 - \sqrt{\frac{s}{D}}l + \frac{1}{2}\frac{s}{D}l^2 - \frac{1}{6}\frac{s^{\frac{3}{2}}}{D^{\frac{3}{2}}}l^3 + \frac{1}{24}\frac{s^2}{D^2}l^4 + O\left(\sqrt{\frac{s}{D}}l\right)$$

把它们带入式（7.18）和式（7.19），忽略高阶项，得到

$$\begin{aligned} p^{\text{tpb}}(s) &= G_{\text{Jp}} J^r(s) + G_{\text{pp}} p^b(s) \\ J^s(s) &= G_{\text{JJ}} J^r(s) + G_{\text{pJ}} p^b(s) \end{aligned} \tag{7.20}$$

式中

$$G_{\text{Jp}} = \frac{-\frac{l}{D} - \frac{l^3}{6D^2}s}{1 + \frac{l^2}{2D}s + \frac{l^4}{24D^2}s^2} \frac{RT}{A}$$

$$G_{\mathrm{pp}}=\frac{1}{1+\dfrac{l^2}{2D}s+\dfrac{l^4}{24D^2}s^2}$$

$$G_{\mathrm{JJ}}=\frac{1}{1+\dfrac{l^2}{2D}s+\dfrac{l^4}{24D^2}s^2}$$

$$G_{\mathrm{pJ}}=\frac{ls}{1+\dfrac{l^2}{2D}s+\dfrac{l^4}{24D^2}s^2}\frac{A}{RT}$$

在式（7.20）中，p^{tpb}代表三相反应界面附近的局部分压，p^b 是气体膨胀中的分压，J^s 是多孔材料表面的气流量，J^r 是气体消耗量或产水量，l 是层的厚度，A 是单位面积，D 是有效扩散系数，R 是气体常数，T 是温度。

式（7.20）正是需要得到的关系式。它可以表明，反应物消耗率 J^r 和局部分压 p^b，是如何确定反应位点处的不可测的分压 p^{tpb}，以及电极表面的质量通量 J^s 的。反应物的消耗速率和局部分压是很容易测到的。

因此，三相反应界面的动态分压以及电极表面的气流量，可以通过气体消耗率和局部分压来确定，并不依赖于沿扩散路径的浓度分布剖面图。

在动态描述中，只与两个参数有关：第一个是扩散层的厚度 l；第二个是扩散系数 D，可以由相关方程计算出来。

7.3.3　扩散动态模型

前面所讲的传递函数可以很容易地转化为常微分方程，常微分方程通常用于模拟和控制。

下面这个模型描述了流入阳极的氢气流：

$$\ddot{J}^s_{\mathrm{H_2}}=-h_1J^s_{\mathrm{H_2}}-h_2\dot{J}^s_{\mathrm{H_2}}+h_1J^r_{\mathrm{H_2}}+h_3\frac{A}{RT}\dot{p}^b_{\mathrm{H_2}} \tag{7.21}$$

三相反应界面附近的氢气分压为

$$\ddot{p}^{\mathrm{tpb}}_{\mathrm{H_2}}=-h_1p^{\mathrm{tpb}}_{\mathrm{H_2}}-h_2\dot{p}^{\mathrm{tpb}}_{\mathrm{H_2}}-h_4\frac{RT}{A}J^r_{\mathrm{H_2}}-\frac{4}{l_{\mathrm{a}}}\frac{RT}{A}\dot{J}^r_{\mathrm{H_2}}+h_1p^b_{\mathrm{H_2}} \tag{7.22}$$

式中，

$$h_1=\frac{24D^2_{\mathrm{H_2}}}{l^4_{\mathrm{a}}},\ h_2=\frac{12D_{\mathrm{H_2}}}{l^2_{\mathrm{a}}},\ h_3=\frac{24D^2_{\mathrm{H_2}}}{l^3_{\mathrm{a}}},\ h_4=\frac{24D_{\mathrm{H_2}}}{l^3_{\mathrm{a}}}$$

流入阴极的氧气通量为

$$\ddot{J}^s_{\mathrm{O_2}}=-o_1J^s_{\mathrm{O_2}}-o_2\dot{J}^s_{\mathrm{O_2}}+o_1J^r_{\mathrm{O_2}}+o_3\frac{A}{RT}\dot{p}^b_{\mathrm{O_2}} \tag{7.23}$$

三相反应界面附近的氧分压为

$$\ddot{p}^{\mathrm{tpb}}_{\mathrm{O_2}}=-o_1p^{\mathrm{tpb}}_{\mathrm{O_2}}-o_2\dot{p}^{\mathrm{tpb}}_{\mathrm{O_2}}-o_4\frac{RT}{A}J^r_{\mathrm{O_2}}-\frac{4}{l_{\mathrm{c}}}\frac{RT}{A}\dot{J}^r_{\mathrm{O_2}}+o_1p^b_{\mathrm{O_2}} \tag{7.24}$$

式中

$$o_1 = \frac{24D_{O_2}^2}{l_c^4},\ o_2 = \frac{12D_{O_2}}{l_c^2},\ o_3 = \frac{24D_{O_2}^2}{l_c^3},\ o_4 = \frac{24D_{O_2}}{l_c^3}$$

在阳极会产生水蒸气，并且会流回燃料流里。水蒸气的动态方程为

$$\ddot{J}_{H_2O}^s = -w_1 J_{H_2O}^s - w_2 \dot{J}_{H_2O}^s + w_1 J_{H_2O}^r + w_3 \frac{A}{RT}\dot{p}_{H_2O}^b \tag{7.25}$$

阳极附近的水蒸气分压为

$$\ddot{p}_{H_2O}^{tpb} = -w_1 p_{H_2O}^{tpb} - w_2 \dot{p}_{H_2O}^{tpb} - w_4 \frac{RT}{A} J_{H_2O}^r - \frac{4}{l_a}\frac{RT}{A}\dot{J}_{H_2O}^r + w_1 p_{H_2O}^b \tag{7.26}$$

式中

$$w_1 = \frac{24D_{H_2O}^2}{l_a^4},\ w_2 = \frac{12D_{H_2O}}{l_a^2},\ w_3 = \frac{24D_{H_2O}^2}{l_a^3},\ w_4 = \frac{24D_{H_2O}}{l_a^3}$$

在这些模型中，为了确定 p^{tpb} 和 J^s，必须知道 p^b 和 J^r 的一阶导数，但是 p^b 和 J^r 都是参数不能直接测定的物理量。

通过引入一个物理变量 ν，输入变量的一阶导数可以近似表示为

$$sU(s) \approx K\left(1 - \frac{1}{\frac{1}{K}s + 1}\right)U(s) \tag{7.27}$$

或者用微分方程的形式：

$$\dot{u} = Ku - \nu$$

$$\dot{\nu} = K^2 u - K\nu \tag{7.28}$$

式中，K 是一个大于 10 的近似因子。

利用这个方法，可以在上述模型中，用可测的变量，求出输入变量的相关变量。

7.3.4　扩散系数

扩散系数可以表示扩散速度的快慢。在燃料电池里，如果有两种不同的材料互相扩散到另一种材料里，此时的系数称为二进制扩散系数。

二进制扩散系数 D_{12} 可以由 Fuller 相关（Welty 等人，1984）来定义：

$$D_{12} = \frac{1.013 \times 10^{-2} T^{1.75}\left(\frac{1}{M_1} + \frac{1}{M_2}\right)^{0.5}}{P[(\sum\nu_i)_1^{1/3} + (\sum\nu_i)_2^{1/3}]^2} \tag{7.29}$$

式中，T 是温度；M_1 和 M_2 分别是气体 1 和气体 2 的摩尔质量；$(\sum\nu_i)_1$ 和 $(\sum\nu_i)_2$ 分别是气体 1 和气体 2 的扩散体积；P 是总压力。

Solheim（1992）曾经报道了 Fuller 相关和测量值之间的关系。

由于燃料电池电极的孔径很小，所以气体分子会频繁地与孔壁碰撞。因此，就需要考虑 Knudsen 扩散机制。总扩散系数为

$$\frac{1}{D}=\frac{1}{D_{12}}+\frac{1}{D_{\mathrm{k}}} \tag{7.30}$$

式中，D_{12}是二进制扩散系数；D_{k} 是 Knudsen 扩散系数。

如果孔径足够大，Knudsen 扩散可以忽略不计。

有效扩散系数受孔隙率和材料弯曲度因素的影响，如下式所示：

$$D_{\mathrm{eff}}=\frac{\varepsilon}{\tau}D \tag{7.31}$$

式中，ε 是孔隙率；τ 是多孔材料的弯曲度；D 是总扩散系数。

在动态模型中，用到的是有效扩散系数。

7.4 燃料输送过程

正如前面所讨论的，扩散过程受两个因素的影响：一个是反应速率，这在前面已讲过，它由外部电流需求和内部负载需求所决定；另一个是在主要反应渠道中各种反应物的分压，它受多种因素的影响，如质量转移、流体输送和热交换。下面将模拟一下燃料和空气的输送渠道，从燃料渠道开始，如图 7.5 所示。

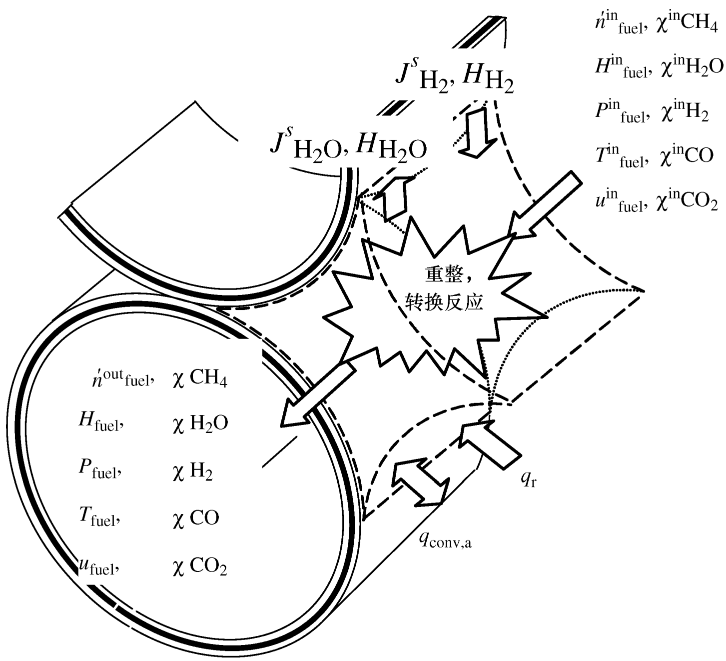


图 7.5　燃料流量部分的质量和能量守恒

7.4.1 重整/转移反应

燃料输送是一个复杂的过程。SOFC 所用的燃料一般是 CH_4、水蒸气、H_2、CO

和 CO_2 的混合物。当这些混合物输送到燃料电池渠道的时候，会发生重整和转移两个反应，从而由 CH_4 和水蒸气产生 H_2 和 CO。

$$\text{重整}: CH_4 + H_2O \rightleftharpoons 3H_2 + CO$$
$$\text{转移}: CO + H_2O \rightleftharpoons H_2 + CO_2 \tag{7.32}$$

两个反应对燃料流量通道里各种组分的浓度影响很大，所以需要研究。要做到这一点，首先需要知道反应速率和平衡常数。反应平衡的定义和解释在第 2 章中已讲到过。

通过 CH_4 消耗速率，由实验定义并确定了转换速率：

$$r_r = -r_{CH_4} = K_r p_{CH_4}^{\alpha} p_{H_2O}^{\beta} \exp\left(-\frac{E_r}{RT_{fuel}}\right) \tag{7.33}$$

式中，K_r 是速率系数；E_r 是活化能；α 和 β 分别是 CH_4 和水蒸气的级数。

与重整反应相比，水-气转换反应更快。它可以很快达到平衡，以至于它的瞬态动态都观察不到。所以在大多数的研究中，只关注平衡常数。转换反应的平衡常数与温度有关，是燃料温度的一个指数函数，定义如下：

$$K_{eq} = \frac{p_{CO_2} p_{H_2}}{p_{CO} p_{H_2O}} = \exp\left(\frac{4276}{T_{fuel}} - 3.961\right) \tag{7.34}$$

根据反应速率的定义，这种转移反应速率方程可以写成

$$r_s = K_s^+ p_{CO} p_{H_2O} - K_s^- p_{CO_2} p_{H_2} \tag{7.35}$$

式中，K_s^+ 和 K_s^- 分别是正向和逆向反应速率常数。

当反应达到平衡时，正向和逆向反应速率是相同的，$r_s = 0$，所以得到

$$K_{eq} = \frac{K_s^+}{K_s^-} \tag{7.36}$$

因此反应速率可以推导为

$$r_s = K_s^- K_{eq} p_{CO} p_{H_2O} - K_s^- p_{CO_2} p_{H_2} \tag{7.37}$$

因为转移反应比其他反应快得多，可以给反应速率系数 K_s 定一个较大的值。将它与已知的平衡常数结合，转移反应速率可以写成

$$r_s = K_s (K_{eq} p_{CO} p_{H_2O} - p_{CO_2} p_{H_2}) \tag{7.38}$$

选择 $K_s = 100$，数字仿真显示转移反应速率几乎是转换反应速率的 10 倍。在式（7.38）中，用 100 取代 K_s，水-气转换速率的模型可以用于写质量平衡方程。

7.4.2　传质过程

燃料流是 CH_4、水蒸气、H_2、CO 和 CO_2 的混合物，其中一些是直接参与电化学反应的，一些不直接参与但是通过反应影响其他反应物。很有必要做一下混合物中各种物质的动态模拟。

在本节中，主要研究各物质的分压。针对式（7.20），本节的重点是确定 p^b，通过质量守恒定律，可以得到分压。压力是表述一定量的分子的一个参数，它们有如下的关系：

$$P = N\frac{RT}{V} \tag{7.39}$$

式中，P 是压力；N 是气体的摩尔数，它代表了气体量；V 是气体的体积；T 是温度；R 是气体常数。

为了方便起见，主要用摩尔数来模拟传质动态。对于混合物，有些参数是很容易测定的，如总压和温度。分压不用直接测定，可以通过公式 $p_i = P\chi_i$ 算出摩尔分数。一种物质的摩尔分数是指该物质的摩尔数占混合物的总摩尔数的比率。

基于摩尔数，将质量守恒定律应用到燃料量上，各物质的连续性方程可以写为

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \dot{n}_i^{\mathrm{in}} - \dot{n}_i^{\mathrm{out}} + \dot{n}_i^{\mathrm{gen}} \tag{7.40}$$

式中，$\dot{n}_i^{\mathrm{in}}$ 和 $\dot{n}_i^{\mathrm{out}}$ 分别代表第 i 种物质进入和流出控制设备的流量；$\dot{n}_i^{\mathrm{gen}}$ 是控制设备中第 i 种物质摩尔数的变化速率。

物质沿着主要流量轴流入控制设备的速率可以用显示气体参数来表示：

$$\dot{n}_i^{\mathrm{in}} = A_{\mathrm{a}}\frac{P_{\mathrm{fuel}}^{\mathrm{in}}}{RT_{\mathrm{fuel}}^{\mathrm{in}}}u_{\mathrm{fuel}}^{\mathrm{in}}\chi_i^{\mathrm{in}} \tag{7.41}$$

式中，A_{a} 是阳极燃料通道的横截面积；$P_{\mathrm{fuel}}^{\mathrm{in}}$、$T_{\mathrm{fuel}}^{\mathrm{in}}$ 和 $u_{\mathrm{fuel}}^{\mathrm{in}}$ 分别是进口压力、温度和燃料流体的速率；χ_i 是第 i 种物质的摩尔分数。

物质沿着主要流量轴的流出速率是

$$\dot{n}_i^{\mathrm{out}} = A_{\mathrm{a}}\frac{N_i}{dV_{\mathrm{a}}}u_{\mathrm{fuel}} \tag{7.42}$$

式中，N_i 是第 i 种物质的摩尔数；dV_{a} 是阳极侧控制量。

除了内部反应的消耗或产生，物质的变化速率还包括其通过正交方向进入主要流向时所产生的扩散，比如 $J_{\mathrm{H_2}}^s$、$J_{\mathrm{O_2}}^s$ 和 $J_{\mathrm{H_2O}}^s$，这些都包含在电化学反应中。

基于以上讨论，CH_4 的积累速率可以写为

$$\frac{\mathrm{d}N_{\mathrm{CH_4}}}{\mathrm{d}t} = A_{\mathrm{a}}\frac{P_{\mathrm{fuel}}^{\mathrm{in}}}{RT_{\mathrm{fuel}}^{\mathrm{in}}}u_{\mathrm{fuel}}^{\mathrm{in}}\chi_{\mathrm{CH_4}}^{\mathrm{in}} - A_{\mathrm{a}}\frac{N_{\mathrm{CH_4}}}{dV_{\mathrm{a}}}u_{\mathrm{fuel}} - Ar_{\mathrm{r}} \tag{7.43}$$

式中，Ar_{r} 是通过反应，甲烷的消耗速率，如式（7.33）所示。

考虑到燃料流中的水蒸气成分，在燃料气流方向会有水蒸气的流入和流出。由于转移或转换反应会有水蒸气的消耗，SOFC 的电化学反应会产生水蒸气，水蒸气以 $J_{\mathrm{H_2O}}^s$ 速率扩散到控制流中。因此，动态模型如下：

$$\frac{\mathrm{d}N_{\mathrm{H_2O}}}{\mathrm{d}t} = A_{\mathrm{a}}\frac{P_{\mathrm{fuel}}^{\mathrm{in}}}{RT_{\mathrm{fuel}}^{\mathrm{in}}}u_{\mathrm{fuel}}^{\mathrm{in}}\chi_{\mathrm{H_2O}}^{\mathrm{in}} - A_{\mathrm{a}}\frac{N_{\mathrm{H_2O}}}{dV_{\mathrm{a}}}u_{\mathrm{fuel}} - J_{\mathrm{H_2O}}^s - Ar_{\mathrm{r}} - r_{\mathrm{s}} \tag{7.44}$$

式中，$-J_{\mathrm{H_2O}}^s$ 是离开阳极表面，进入控制流的水蒸气；r_{s} 是通过转移反应水蒸气的消耗速率。

燃料流向上，转移或转换反应分别以 $3r_{\mathrm{s}}$ 和 r_{s} 的速率产生 H_2。电化学反应会消耗 H_2 来产生电力：

$$\frac{dN_{H_2}}{dt}=A_a\frac{P_{fuel}^{in}}{RT_{fuel}^{in}}u_{fuel}^{in}\chi_{H_2}^{in}-A_a\frac{N_{H_2}}{dV_a}u_{fuel}-J_{H_2}^{s}+3Ar_r+r_s \tag{7.45}$$

重整反应也会产生 CO，一部分 CO 会由转移反应消耗掉。它也可能扩散到三相反应界面，并参与电化学反应。然而，在三相反应界面，H_2 和 CO 竞相参与电化学反应的机理还不明确，CO 扩散到电极的速率也不清楚。如果不考虑 CO 的电化学反应，方程可以写为

$$\frac{dN_{CO}}{dt}=A_a\frac{P_{fuel}^{in}}{RT_{fuel}^{in}}u_{fuel}^{in}\chi_{CO}^{in}-A_a\frac{N_{CO}}{dV_a}u_{fuel}+Ar_r-r_s \tag{7.46}$$

CO_2 的产生是相对简单的，只有在转移反应中燃气流向上时会产生 CO_2：

$$\frac{dN_{CO_2}}{dt}=A_a\frac{P_{fuel}^{in}}{RT_{fuel}^{in}}u_{fuel}^{in}\chi_{CO_2}^{in}-A_a\frac{N_{CO_2}}{dV_a}u_{fuel}+r_s \tag{7.47}$$

如果确实要考虑电化学反应中的 CO，可以对基于成分的摩尔通量做合理的假设：

$$\begin{aligned}J_{CO}^{s}&=\frac{\chi_{CO}}{\chi_{CO}+\chi_{H_2}}J^{s}\\ J_{H_2}^{s}&=\frac{\chi_{H_2}}{\chi_{CO}+\chi_{H_2}}J^{s}\end{aligned} \tag{7.48}$$

式中，J^s 是输出电流 I 所需要的反应物流量，根据这种假设，H_2、CO 和 CO_2 的传质过程模型可以修改为包括电化学反应中的 CO。

7.4.3　动量传递

如前所述，如果想知道物质的摩尔数和摩尔分数，必须知道燃料的流速。通过反应，燃料流的组分会发生变化，如从 H_2 到 H_2O，从 CO 到 CO_2，沿着流动方向，燃料流的密度也会发生变化。所以，燃料流速也会发生变化。

燃料流速决定了燃料留在通道里的时间，这样也就影响了燃料的使用。它还会影响热传导系数［之后将在式（7.67）中介绍］，反过来，影响电池的温度。更重要的是，流速可以直接被控制，可以看做是控制 SOFC 性能的一个变量。所以，知道速度动态是有价值的，将动量守恒定律运用到运动方程中，速度可以这样定义，如第 2 章所示：

$$\frac{d(\sum(N_iM_i)u_{fuel})}{dt}=\sum\dot{m}_i^{in}u_{fuel}^{in}-\sum\dot{m}_i^{out}u_{fuel}^{out}-J_{H_2}^{s}M_{H_2}u_{fuel}+\sum F_a \tag{7.49}$$

式中

$$\sum\dot{m}_i^{in}u_{fuel}^{in}=A_a\frac{P_{fuel}^{in}}{RT_{fuel}^{in}}\sum(\chi_i^{in}M_i)u_{fuel}^{in2}$$

$$\sum\dot{m}_i^{out}u_{fuel}^{out}=A_a\frac{\sum(N_iM_i)}{dV_a}u_{fuel}^{2}$$

式（7.49）右边的前两项分别是在主流方向上流入和流出的动量速率。第三

项是流出控制流量沿着正交氢气流的动量速率。电化学反应会产生 H_2O，H_2O 会扩散到控制流中，假设 H_2O 的轴向速度是零。因此，从正交方向到主流向的动量应该是零。最后一项代表作用于燃料流的力的总和。

在燃料的流动方向上有两种力作用于燃料流：一是上游压力梯度，它是燃料流移动的动力；二是沿着通道壁的剪应力，它在相反方向上有影响。所以作用于燃料流的力的总和为

$$\sum F_{\mathrm{a}} = A_{\mathrm{a}} \Delta P_{\mathrm{a}} - A\tau_{\mathrm{w,a}} \tag{7.50}$$

压力梯度 ΔP_{a} 可以由 Poiseuille 定律确定，剪应力 τ_{w} 可以由摩擦系数 f 计算出来：

$$f = \frac{\tau_{\mathrm{w}}}{\frac{1}{2}\rho u^2} \tag{7.51}$$

系数 fRe 是特殊通道的常数，适用于流动是层流型时，它的具体值可以由 Perry 等人（1997a）的相关文献中查到。对于一个动荡的燃料流，表面粗糙度影响着 f 的值，可以由表面粗糙度和雷诺数求出来。

当在有限的体积模型里考虑无边界条件时，假定作用于燃料流的力是平衡的，即 $\sum F_{\mathrm{a}} = 0$。

7.4.4 能量转移和热交换

温度几乎影响其他所有的变量，它由多个因素决定。例如，燃气温度受多重因素的影响、转换反应热、流体带来的内能还有周围环境中的热传递。可以合理地假设，只有很少量的动能转化成内能。所设计的燃料电池不是在高流速下运行，因此可以忽略转换。

在编写气相能量平衡中，使用了焓的概念，如第 2 章所述，焓不仅包括内能，还包括体积功。因此，反应热可以简单地只考虑参加反映的反应物焓变。

燃料流量的温度变化主要受反应热和热交换的影响。考虑到燃料流控制体积的能量平衡，温度动态可以像第 2 章中那样推导：

$$\begin{aligned}\frac{\mathrm{d}\left(\sum (N_i C_{v,i}) T_{\mathrm{fuel}}\right)}{\mathrm{d}t} &= \sum (\dot{n}_i^{\mathrm{in}} H_i^{\mathrm{in}}) - \sum (\dot{n}_i^{\mathrm{out}} H_i) \\ &\quad - J_{\mathrm{H_2}}^{s} H_{\mathrm{H_2}} - J_{\mathrm{H_2O}}^{s} H_{\mathrm{H_2O}} - h_{\mathrm{a}} A (T_{\mathrm{fuel}} - T_{\mathrm{cell}}) + q_{\mathrm{ref}}\end{aligned} \tag{7.52}$$

式中

$$\sum (\dot{n}_i^{\mathrm{in}} H_i^{\mathrm{in}}) = A_{\mathrm{a}} \frac{P_{\mathrm{fuel}}^{\mathrm{in}}}{RT_{\mathrm{fuel}}^{\mathrm{in}}} u_{\mathrm{fuel}}^{\mathrm{in}} \sum (\chi_i^{\mathrm{in}} H_i^{\mathrm{in}})$$

$$\sum (\dot{n}_i^{\mathrm{out}} H_i) = A_{\mathrm{a}} \frac{\sum (N_i H_i)}{dV_{\mathrm{a}}} u_{\mathrm{fuel}}$$

式（7.52）右边的前四项是燃气流入和流出时所携带的能量，不包含参与转换反应的那部分燃料流。第五项是当燃料通过电池管道壁的时候产生的对流传热。

最后一项代表由于转换反应产生的热量。如 Qi 等人（2008）的具体阐述，转换反应可能被认为发生在阳极表面，因为阳极的 Ni 元素可以作为反应的催化剂。因此，这一项可以表示从反应物到生成物的净焓变。可以表示如下：

$$q_{\mathrm{ref}} = A r_{\mathrm{r}} \Delta H | T_{\mathrm{fuel}} \tag{7.53}$$

式中，r_{r} 是转换反应速率；$\Delta H = (3H_{H_2} + H_{CO} - H_{CH_4} - H_{H_2O})$ 是反应热。

式（7.52）右边的前四项没有直接包括水-气转换反应的热量，相反，它已经隐含在这四项里面了。因为转换反应是发生在燃料流内部的，焓变已经包含了反应带来的能量变化。在前一节中已经确定了各物质的流量。转换反应是放热的，这就导致燃料内能的增加。

7.5　空气输送过程

空气被吹到阴极侧，作为氧气的来源。没有外部反应伴随燃料流。但是，氧气的消耗会影响整个燃料电池的动态系统，需要定义一下流过通道的气流。

如图 7.6 所示，空气流经两个部分：从入口到电池管的底部，然后在管底部转动起来，通过阴极通道向相反的方向流动。第一部分是一个简单的引导部分，没有反应，没有传质。只需要确定第二部分空气输送质量的模型。

7.5.1　阴极通道传质

研究阴极通道里传质动态的控制体积不包括阴极侧通道的气流（见图 7.6）。

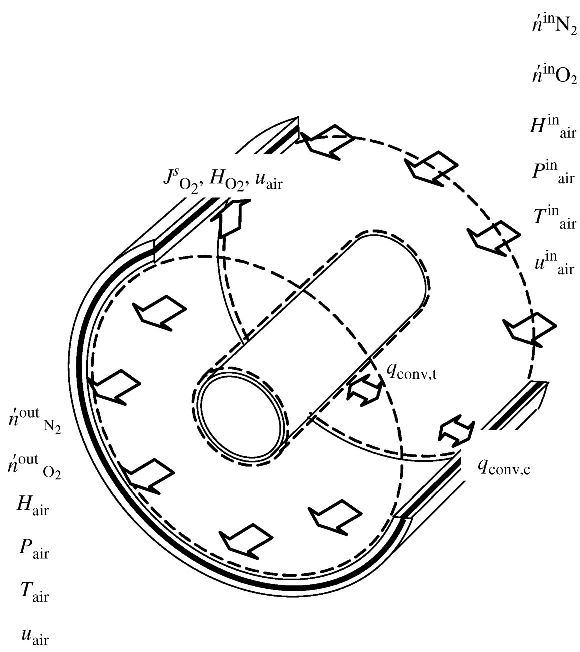


图 7.6　气流部分的质量和能量守恒

在这个控制体积里，O_2 离开气流扩散到阴极，在三相反应界面参与电化学反应，在这里，O_2 与电子结合，被还原成 O^{2-}。然后这些阳离子通过电解液到达阳极侧的三相反应界面，参与电化学反应。不会有反应产物扩散回阴极侧的气流中。

可以合理地忽略气流中其他次要的成分，假设气流中只包括 N_2 和 O_2，N_2 不参与电化学反应，它的变化率只受气流的影响，可以定义如下：

$$\frac{\mathrm{d}N_{N_2}}{\mathrm{d}t}=A_c\frac{P_{air}^{in}}{RT_{air}^{in}}(1-\chi_{O_2}^{in})u_{air}^{in}-A_c\frac{N_{N_2}}{dV_c}u_{air} \tag{7.54}$$

式中，A_c 是阴极通道的横截面积；dV_c 是体积；$(1-\chi_{O_2}^{in})$ 是 N_2 的进气摩尔分数。

O_2 的动态质量平衡方程可以定义如下：

$$\frac{\mathrm{d}N_{O_2}}{\mathrm{d}t}=A_c\frac{P_{air}^{in}}{RT_{air}^{in}}\chi_{O_2}^{in}u_{air}^{in}-A_c\frac{N_{O_2}}{dV_c}u_{air}-J_{O_2}^{s} \tag{7.55}$$

式中，$J_{O_2}^{s}$代表电化学反应中消耗的 O_2 成分。

7.5.2 阴极通道动量传递

与阳极一侧相似，使用与前一节中提出的相同的方法，阴极侧流速的动态方程如下：

$$\frac{\mathrm{d}(\sum(N_jM_j)u_{air})}{\mathrm{d}t}=A_c\frac{P_{air}^{in}}{RT_{air}^{in}}\sum(\chi_j^{in}M_j^{in})u_{air}^{in2}-A_c\frac{\sum(N_jM_j)}{dV_c}u_{air}^2-J_{O_2}^{s}M_{O_2}u_{air}+\sum F_c \tag{7.56}$$

式中，M_j 是 N_2 或 O_2 的摩尔质量；$\sum F_c=A_c\Delta P_c-(A_1+A_2)\tau_{w,a}$是作用于气流的力的总和。

7.5.3 阴极通道能量传递

空气与电池管和注入管进行热交换，因此有两个热对流过程。空气的加热和能量传递过程可以用如下方程表示：

$$\begin{aligned}\frac{\mathrm{d}(\sum(N_jC_{v,j})T_{air})}{\mathrm{d}t}=&A_c\frac{P_{air}^{in}}{RT_{air}^{in}}u_{air}^{in}H_{air}^{in}-A_c\frac{\sum N_jH_{air}}{dV_c}u_{air}\\&-J_{O_2}^{s}H_{O_2}-h_cA_2(T_{air}-T_{cell})-h_cA_1(T_{air}-T_{tube})\end{aligned} \tag{7.57}$$

式中，h_c 是阴极侧的热传递系数；A_1 和 A_2 分别是电池管的内表面积和注入管的外表面积。

N_2 和 O_2 没有单独出现，因为空气的焓这一项会自动根据组分而改变。

7.5.4 注入通道空气

注入通道里的空气流动过程如图 7.7 所示。这个通道的气流在热交换过程中起着重要的作用。在这个控制设备中，不会有反应和传质过程发生。但是，气流可以加热或冷却注入管，从而影响整个电池温度。

考虑控制设备中的能量守恒，温度动态可以通过能量变化的方程定义如下：

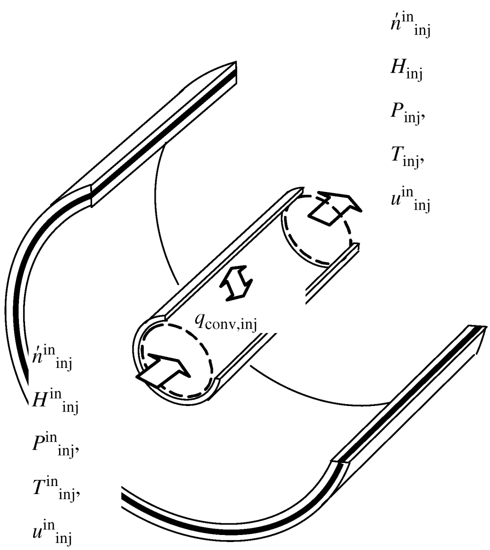


图 7.7　气流在注入管中的质量和能量守恒

$$N_{air}C_{v,air}\frac{dT_{inj}}{dt}=A_t\frac{P^{in}_{inj}}{RT^{in}_{inj}}(H^{in}_{inj}-H_{inj})u^{in}_{air}-h_tA_0(T_{inj}-T_{tube}) \tag{7.58}$$

式中，A_t 是注入管通道的内部横截面积。

7.6　SOFC 温度

针对前面部分提出的不同的动态方程，可以注意到，在目前所研究的所有的动态方程中，包括电化学反应、气流的热传递，转换和转移反应等，温度都会是一个常见的因素。为了描述温度，需要应用第 2 章中所介绍的热力学和热传递原理。

所有 3 个传热机制，以及由于反应产生的能量变化，在 SOFC 中都会有（见图 7.8）。SOFC 的所有组成部分都参与热交换，并且影响电池的温度。

7.6.1　动态能量交换过程

通过电化学反应，Gibbs 自由能转化为电能，以释放反应热的方式维持焓平衡。大部分电能释放到外部负载上，SOFC 的内阻也会消耗掉一部分能量。由于电池和周围环境存在温度差异，所以会通过强制对流、辐射和传导的方式发生热交换。

根据热力学第一定律，控制设备中内能积累的速率方程可以这样写：

$$m_{cell}C_{p,cell}\frac{dT}{dt}=\sum J^s_iH_i-E_e-q_{conv,a}-q_{conv,c}-q_{rad}-q_{ref} \tag{7.59}$$

式中，$\sum J^s_iH_i$ 是焓变的总和；E_e 是供应到外部负载的电力；$q_{conv,a}$ 和 $q_{conv,c}$ 分别是通

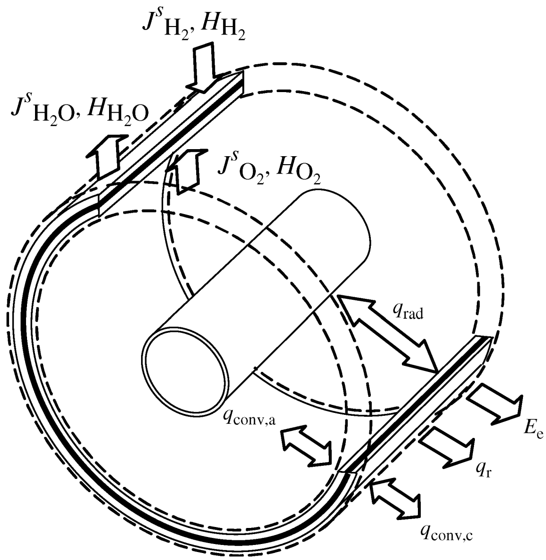


图 7.8　SOFC 管中的能量交换

过热对流与阳极和阴极发生热交换的热气流；q_{rad}是热交换和热辐射的热流；q_{ref}是转换反应产生的热（这在 7.4.4 节中讲到过）。

7.6.1.1　焓

焓或生成焓是一种更好的用来描述气体能量的概念。它是物质内能与体积和压力乘积的总和，在本节中，焓广泛地用于能量动态中。

温度对生成焓有显著的影响。SOFC 中，物质的温度和焓之间的关系可以用一个有关温度的多项式来表示，在附录中可以找到。

如果只有氢参与阳极电化学反应，生成焓变可以这样定义：

$$\sum J_i^s H_i = J_{H_2}^s H_{H_2}\big|_{T_{fuel}} + J_{O_2}^s H_{O_2}\big|_{T_{air}} + J_{H_2O}^s H_{H_2O}\big|_{T_{cell}} \tag{7.60}$$

式中，$J_{H_2}^s$、$J_{O_2}^s$和 $J_{H_2O}^s$是流入控制设备的流通率。

7.6.1.2　电力

SOFC 的电力输出是外部负载的唯一的能量来源：

$$E_e = V_{out} I = \frac{V_{out}^2}{R_{load}} \tag{7.61}$$

在接下来的三节中，会有对热交换过程，包括热传导、对流和辐射的详细研究。

7.6.2　热传导

热传导的概念已经在第 2 章中讲到过，SOFC 的热传导主要发生在固相上。沿轴向的热传导可以这样定义：

$$q_{cond} = -k_{cell} A_{cross} \frac{dT_{cell}}{dz} \tag{7.62}$$

式中，k_{cell}是电池片的导热系数；A_{cross}是横截面积；T_{cell}是温度。

对于管式的 SOFC，热传导由沿着轴向的 $k_{cell} * A_{cross}$决定，远小于径向的强制对流和辐射传热。因此 q_{cond}很小，可以忽略。

沿径向方向电池非常薄，温差很小，因此可以合理地假设沿径向方向没有温度变化。

7.6.3　对流

当燃料和气流经 SOFC 的表面时，强制对流是一个重要的传热方式，它使得电池管和气流之间发生热交换：

$$q_{conv} = hA(T_w - T_f) \tag{7.63}$$

式中，h 是热传导系数；A 是 SOFC 的面积；T_w 是固体壁的温度；T_f 是流体的温度。

从第 2 章已经了解到，对流是几个不同机制的组合，因此热传导系数 h 受到一些流体性质的影响。可以通过一个经验公式，由无量纲的努赛尔数 Nu、雷诺数 Re 和普朗克数 Pr 确定对流传热系数。文献中经常采用这几个无量纲数来描述流体性质。

努赛尔数定义如下：

$$Nu = \frac{hD_f}{k_f} \tag{7.64}$$

式中，D_f 是流体直径；k_f 是流体的热导率；h 是传热系数。

那么，传热系数可以表示为

$$h = Nu\frac{k_f}{D_f} \tag{7.65}$$

努赛尔数 Nu 可以表示为雷诺数 Re 和普朗克数 Pr 的函数（Welty 等人，1984）：

$$\begin{aligned}&\text{层流:}Nu = 0.332Pr^{1/3}Re^{1/2}\\&\text{紊乱流:}Nu = 0.023Pr^{1/3}Re^{0.8}\end{aligned} \tag{7.66}$$

对于所研究的 SOFC，阳极和阴极侧流的雷诺数可以分别定为 66 和 287，它们都属于层流式，我们可以利用层流的相关性。

对于层流式流体，可以利用用雷诺数 Re 和普朗克数 Pr 的相关性，传热系数可以确定为

$$h = 0.332\frac{k_f}{D_f}\left(\frac{C_{v,f}\mu_f}{k_f}\right)^{1/3}\left(\frac{D_fM_f}{\mu_f}\right)^{1/2}\left(\frac{u_fP_f}{RT_f}\right)^{1/2} \tag{7.67}$$

因此，强制对流热传递可以表示为

$$\begin{aligned}q_{conv,a} &= f_a\left(\frac{u_{fuel}P_{fuel}}{RT_{fuel}}\right)^{1/2}A(T_{cell} - T_{fuel})\\q_{conv,c} &= f_c\left(\frac{u_{air}P_{air}}{RT_{air}}\right)^{1/2}A_2(T_{cell} - T_{air})\end{aligned} \tag{7.68}$$

式中

$$f_a = 0.332\frac{k_{fuel}}{D_a}\left(\frac{C_{v,fuel}\mu_{fuel}}{k_{fuel}}\right)^{1/3}\left(\frac{D_a M_{fuel}}{\mu_{fuel}}\right)^{1/2}$$

$$f_c = 0.332\frac{k_{air}}{D_c}\left(\frac{C_{v,air}\mu_{air}}{k_{air}}\right)^{1/3}\left(\frac{D_c M_{air}}{\mu_{air}}\right)^{1/2}$$

随着反应进行，燃料混合物的组成会发生变化，所以组分性能的加权平均能更准确地表示流特性：

$$\begin{aligned} k_{fuel} &= \sum \chi_i k_i \\ C_{v,fuel} &= \sum \chi_i C_{v,i} \\ \mu_{fuel} &= \sum \chi_i \mu_i \\ M_{fuel} &= \sum \chi_i M_i \end{aligned} \tag{7.69}$$

式中，χ 代表摩尔分数；i 代表物质种类，包括 H_2、H_2O、CH_4、CO 和 CO_2。

7.6.4　辐射

燃料电池运行在一个较高的温度下，辐射就成为了一种不可忽视的传热机制。在绝对温度 T 下，辐射通量为

$$q = \varepsilon E_b = \varepsilon \sigma T^4 \tag{7.70}$$

式中，ε 是辐射系数，介于 0 和 1 之间；σ 是 Stefan-Boltzmann 常数。

一个 SOFC 堆栈由多个 SOFC 管组成。每一个管都有相同的尺寸、排布和运行环境。因此可以假定管的表面温度是相同的，管的外表面（阳极侧）的净辐射传热量接近于零。

燃料/空气吸收的辐射热量也可以被认为是微不足道的。

辐射传热主要出现在电池管的内表面和注入管的外表面之间。两者间净辐射热为

$$q_{rad} = \frac{\sigma}{R_{rad}} A_2 (T_{cell}^4 - T_{tube}^4) \tag{7.71}$$

式中，R_{rad}是传热阻力。

类似于电路电阻，它可以这样计算：

$$R_{rad} = \frac{1-\varepsilon_{cell}}{\varepsilon_{cell}} + \frac{1}{F_{c-t}} + \frac{1-\varepsilon_{tube}}{\varepsilon_{tube}}\frac{A_2}{A_1} \tag{7.72}$$

式中，ε_{cell}和 ε_{tube}分别是电池和注入管的辐射率；F_{c-t}是从电池到注入管的视角系数；A_1 和 A_2 分别是注入管的外表面积和电池管的内表面积。

一般从 A_1 面到 A_2 面的视角系数表示为

$$F_{A_1-A_2} = \frac{1}{A_1}\int_{A_1}\int_{A_2}\frac{\cos\theta_1\cos\theta_2}{\pi r^2}dA_2 dA_1 \tag{7.73}$$

式中，A_1 和 A_2 分别是 A_1 面和 A_2 面的面积，其他参数如图 7.9 所示。

对于管式 SOFC，A_1 面到 A_2 面的视角系数表示为

$$F_{c-t}=\frac{r_{t,outer}}{r_{c,inner}}\quad\frac{\text{长度}}{r_{c,inner}}\longrightarrow\infty \tag{7.74}$$

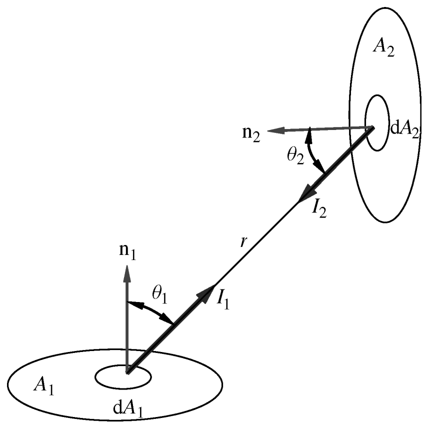


图 7.9　视角系数的定义

7.6.5　电池温度模型

将式（7.60）、式（7.61）、式（7.68）和式（7.71）带入式（7.59），推导出如下方程来描述电池温度：

$$m_{cell}C_{p,cell}\frac{dT}{dt}=J^s_{H_2}H_{H_2}\big|_{T_{fuel}}+J^s_{O_2}H_{O_2}\big|_{T_{air}}+J^s_{H_2O}H_{H_2O}\big|_{T_{cell}}-\frac{V^2_{out}}{R_{load}}-q_{ref}$$
$$-f_a\left(\frac{u_{fuel}P_{fuel}}{RT_{fuel}}\right)^{1/2}A(T_{cell}-T_{fuel})-f_c\left(\frac{u_{air}P_{air}}{RT_{air}}\right)^{1/2}A_2(T_{cell}-T_{air}) \tag{7.75}$$
$$-\frac{\sigma}{R_{rad}}A_2(T^4_{cell}-T^4_{tube})$$

7.6.6　注入管温度模型

作为导管，氧化铝注入管在热传递过程中起着很大的作用，如图 7.10 所示。

管与阴极侧气流及管内的气流发生热交换，管还通过辐射与电池内壁发生热交换。注入管的温度动态模型如下：

$$m_{tube}C_{p,tube}\frac{dT_{tube}}{dt}=\frac{\sigma}{R_{rad}}A_1(T^4_{cell}-T^4_{tube})-h_cA_1(T_{tube}-T_{air})-h_tA_0(T_{tube}-T_{inj}) \tag{7.76}$$

式中，h_t 是注入管内部的对流传热系数，可以表示如下：

$$h_t=f_t\left(\frac{u_{inj}P_{inj}}{RT_{inj}}\right)^{1/2}$$

式中

$$f_t=0.332\,\frac{k_{air}}{D_t}\left(\frac{C_{v,air}\mu_{air}}{k_{air}}\right)^{1/3}\left(\frac{D_tM_{air}}{\mu_{air}}\right)^{1/2}$$

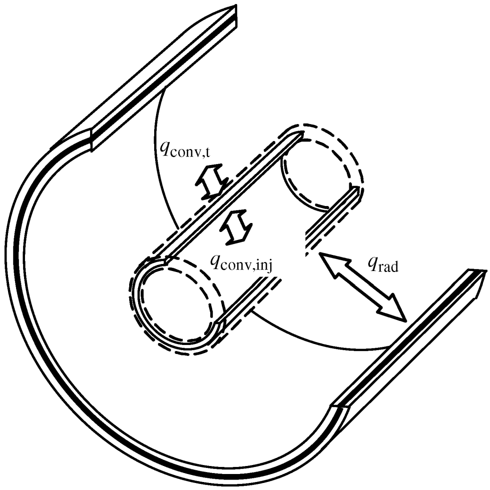


图 7. 10　注入管部分能量守恒

7.7　最终动态模型

以状态空间模型为基础，可以用推导出来的一些微分方程描述 SOFC 的动态行为，这对于控制应用程序很方便，也可以进一步开发。状态空间模型是用一组状态变量的常微分方程描述一个动态的过程。状态变量是一组对于描述过程内部关系相当重要的变量。空间状态模型组成一个著名的描述动态系统的模型，尤其是对于多输入和多输出系统。它广泛应用在建模、控制和仿真中。在本节中，SOFC 的开发动态微分方程将被转换为一个状态空间模型。

7.7.1　I/O 变量

将开发的微分方程转换为状态空间模型，第一步就是确定输入和输出变量。通常，选择可以物理性操作或繁杂的变量作为模型的输入变量，物理输出和感兴趣的状态变量作为模型的输出变量。

第一个输入变量是外部负载阻抗。在正常操作环境下，外部负载决定电流，电流又反过来决定了电化学反应以及质量和热量转移。为了简化解决方案，外部负载可以假定为一个纯电阻 R_{load}。

下一组的输入变量代表入口处的燃料状态，包括压力 P_{fuel}^{in}、温度 T_{fuel}^{in} 和流速 u_{fuel}^{in}。虽然燃料流中每种物质的进气摩尔分数不能操作，但这些变量的变化会影响 SOFC 的性能。因此，由于是障碍，进气摩尔分数也定义为输入变量。

第三组输入变量是阴极侧通道里进气气流性质，包括压力 P_{air}^{in}、温度 T_{air}^{in}、流速 u_{air}^{in} 以及氧气的摩尔分数 χ_{O_2}。

最后一组输入变量是空气注入压力 P_{inj}^{in}、温度 T_{inj}^{in} 和流速 u_{inj}^{in}。

输出变量是电压 V_{out}、电流 I 以及与出口流的物理性质相关的变量，见表 7.1。

尽管电池温度 T_{cell} 是一个状态变量，但它可以很容易地转换为一个输出变量，其他任何状态变量也一样。

表 7.1　输入和输出变量

输入变量	输出变量
R_{load}	V_{out}，I
P_{fuel}^{in}，T_{fuel}^{in}，u_{fuel}^{in}	P_{fuel}，T_{fuel}，u_{air}
$\chi_{CH_4}^{in}$，$\chi_{H_2O}^{in}$，$\chi_{H_2}^{in}$，χ_{CO}^{in}，$\chi_{CO_2}^{in}$	χ_{CH_4}，χ_{H_2O}，χ_{H_2}，χ_{CO}，χ_{CO_2}
P_{air}^{in}，T_{air}^{in}，u_{air}^{in}，$\chi_{O_2}^{in}$	P_{air}，T_{air}，u_{air}，χ_{O_2}
P_{inj}^{in}，T_{inj}^{in}，u_{inj}^{in}	P_{inj}，T_{inj}，u_{inj}

7.7.2　状态空间模型

定义输入矢量 $\boldsymbol{u}$ 和输出矢量 $\boldsymbol{y}$ 如下：

$$\begin{aligned}\boldsymbol{u}=(&R_{load}\quad P_{fuel}^{in}\quad T_{fuel}^{in}\quad u_{fuel}^{in}\quad \chi_{CH_4}^{in}\quad \chi_{H_2O}^{in}\quad \chi_{H_2}^{in}\quad \chi_{CO}^{in}\quad \chi_{CO_2}^{in}\\ &P_{air}^{in}\quad T_{air}^{in}\quad u_{air}^{in}\quad \chi_{O_2}^{in}\quad P_{inj}^{in}\quad T_{inj}^{in}\quad u_{inj}^{in})^{T}\end{aligned}\tag{7.77}$$

$$\begin{aligned}\boldsymbol{y}=(&V_{out}\quad I\quad P_{fuel}\quad T_{fuel}\quad u_{fuel}\quad \chi_{CH_4}\quad \chi_{H_2O}\quad \chi_{H_2}\quad \chi_{CO}\quad \chi_{CO_2}\\ &P_{air}\quad T_{air}\quad u_{air}\quad \chi_{O_2}\quad P_{inj}\quad T_{inj}\quad u_{inj})^{T}\end{aligned}\tag{7.78}$$

定义状态矢量如下：

$$\begin{aligned}\boldsymbol{x}=(&V_{ct}\quad J_{H_2}^{s}\quad \dot{J}_{H_2}^{s}\quad J_{O_2}^{s}\quad \dot{J}_{O_2}^{s}\quad J_{H_2O}^{s}\quad \dot{J}_{H_2O}^{s}\quad p_{H_2}^{tpb}\quad \dot{p}_{H_2}^{tpb}\quad p_{O_2}^{tpb}\\ &\dot{p}_{O_2}^{tpb}\quad p_{H_2O}^{tpb}\quad \dot{p}_{H_2O}^{tpb}\quad \nu_{R}\quad T_{cell}\quad T_{tube}\quad T_{fuel}\quad T_{air}\quad T_{inj}\\ &N_{CH_4}\quad N_{H_2O}\quad N_{H_2}\quad N_{CO}\quad N_{CO_2}\quad N_{N_2}\quad N_{O_2}\quad u_{fuel}\quad u_{air})^{T}\end{aligned}\tag{7.79}$$

将式（7.33）和式（7.38）代入式（7.43）和式（7.47），将式（7.33）和式（7.53）代入式（7.75）和式（7.52），结合 7.2.2 节中所讲的电力变换的动态方程和重新整理的式（7.75）、式（7.43）~式（7.47）、式（7.52）、式（7.49）以及式（7.54）~式（7.58），可以导出如下状态空间模型：

状态：

$$\dot{x}_1=\frac{1}{R_{ct}C_{ct}}E-\frac{1}{R_{ct}C_{ct}}x_1-\frac{1}{C_{ct}}\frac{x_1}{u_1+R_o}$$

$$\dot{x}_2=x_3$$

$$\dot{x}_3=-h_1x_2-h_2x_3+h_1\frac{1}{2F}\frac{x_1}{u_1+R_o}+h_3\frac{A}{dV_a}\left(\dot{x}_{22}+\frac{x_{22}}{x_{17}}\dot{x}_{17}\right)$$

$$\dot{x}_4=x_5$$

$$\dot{x}_5=-o_1x_4-o_2x_5+o_1\frac{1}{4F}\frac{x_1}{u_1+R_o}+o_3\frac{A}{dV_c}\left(\dot{x}_{26}+\frac{x_{26}}{x_{18}}\dot{x}_{18}\right)$$

$$\dot{x}_6=x_7$$

$$\dot{x}_7=-w_1x_6-w_2x_7+w_1\frac{1}{2F}\left(\frac{-x_1}{u_1+R_o}\right)+w_3\frac{A}{dV_a}\left(\dot{x}_{21}+\frac{x_{21}}{x_{17}}\dot{x}_{17}\right)$$

$$\dot{x}_8=x_9$$

$$\dot{x}_9 = -h_1 x_8 - h_2 x_9 - h_4 \frac{Rx_{15}}{A} \frac{1}{2F} \frac{x_1}{u_1 + R_o} + h_1 \frac{R}{dV_a} x_{22} x_{17} - \frac{Rx_{15}}{A} \frac{4}{la} \frac{1}{2F}\left[\frac{\dot{x}_1}{u_1 + R_o} - \frac{x_1}{(u_1 + R_o)^2}(Ku_1 - x_{14})\right]$$

$$\dot{x}_{10} = x_{11}$$

$$\dot{x}_{11} = -o_1 x_{10} - o_2 x_{11} - o_4 \frac{Rx_{15}}{A} \frac{1}{4F} \frac{x_1}{u_1 + R_o} + o_1 \frac{R}{dV_c} x_{26} x_{18} - \frac{Rx_{15}}{A} \frac{4}{lc} \frac{1}{4F}\left[\frac{\dot{x}_1}{u_1 + R_o} - \frac{x_1}{(u_1 + R_o)^2}(Ku_1 - x_{14})\right]$$

$$\dot{x}_{12} = x_{13}$$

$$\dot{x}_{13} = -w_1 x_{12} - w_2 x_{13} - w_4 \frac{Rx_{15}}{A} \frac{1}{2F} \frac{-x_1}{u_1 + R_o} + w_1 \frac{R}{dV_a} x_{21} x_{17} - \frac{Rx_{15}}{A} \frac{4}{la} \frac{1}{2F}\left[-\frac{\dot{x}_1}{u_1 + R_o} + \frac{x_1}{(u_1 + R_o)^2}(Ku_1 - x_{14})\right]$$

$$\dot{x}_{14} = K^2 u_1 - Kx_{14}$$

$$\dot{x}_{15} = \frac{1}{m_{\text{cell}} C_{\text{p,cell}}}\left[x_2 H_{H_2} + x_4 H_{O_2} + x_6 H_{H_2O} - \frac{u_1}{(u_1 + R_o)^2} x_1^2 - f_a\left(x_{27}\frac{\sum_{20}^{24} x_i}{dV_a}\right)^{\frac{1}{2}} A(x_{15} - x_{17}) - f_c\left(x_{28}\frac{x_{25} + x_{26}}{dV_c}\right)^{\frac{1}{2}} A_2(x_{15} - x_{18}) - \frac{\sigma}{R_{\text{rad}}} A_2(x_{15}^4 - x_{16}^4) - AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)\Delta H_r\right]$$

$$\dot{x}_{16} = \frac{1}{m_{\text{tube}} C_{\text{p,tube}}}\left[\frac{\sigma}{R_{\text{rad}}} A_2(x_{15}^4 - x_{16}^4) - f_c\left(x_{28}\frac{x_{25} + x_{26}}{dV_c}\right)^{\frac{1}{2}} A_1(x_{16} - x_{18}) - f_t\left(u_{16}\frac{u_{14}}{Rx_{19}}\right)^{\frac{1}{2}} A_0(x_{16} - x_{19})\right]$$

$$\dot{x}_{17} = \frac{1}{\sum_{20}^{24}(x_i C_{v,i})}\left[A_a \frac{u_2}{Ru_3} u_4 \sum_5^9 (u_i H_i^{\text{in}}) - A_a x_{27}\frac{\sum_{20}^{24}(x_i H_i)}{dV_a} - f_a\left(x_{27}\frac{\sum_{20}^{24} x_i}{dV_a}\right)^{\frac{1}{2}} A(x_{17} - x_{15}) - x_2 H_{H_2} - x_6 H_{H_2O} + AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)\Delta H_r - x_{17}\sum_{20}^{24}(\dot{x}_i C_{v,i})\right]$$

$$\dot{x}_{18} = \frac{1}{\sum_{25}^{26}(x_j C_{v,j})}\left[A_c \frac{u_{10}}{Ru_{11}} u_{12} H_{\text{air}}^{\text{in}} - A_c x_{28}\frac{\sum_{25}^{26} x_j}{dV_c} H_{\text{air}} - x_4 H_{O_2} - f_c\left(x_{28}\frac{\sum_{24}^{26} x_j}{dV_c}\right)^{\frac{1}{2}} (A_2(x_{18} - x_{15}) + A_1(x_{18} - x_{16})) - x_{18}\sum_{25}^{26}(\dot{x}_j C_{v,j})\right]$$

$$\dot{x}_{19} = \frac{1}{\frac{u_{14} dV_t}{Ru_{15}} C_{\text{v,air}}}\left[A_t \frac{u_{14}}{Ru_{15}} u_{16}(H_{\text{air}}^{\text{in}} - H_{\text{air}}) - f_t\left(u_{16}\frac{u_{14}}{Rx_{19}}\right)^{\frac{1}{2}} A_0(x_{19} - x_{16})\right]$$

$$\dot{x}_{20} = A_a \frac{u_2}{Ru_3} u_4 u_5 - A_a x_{27}\frac{x_{20}}{dV_a} - AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)$$

$$\dot{x}_{21} = A_a \frac{u_2}{Ru_3} u_4 u_6 - A_a x_{27}\frac{x_{21}}{dV_a} - AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)$$

$$-AK_s\left[\exp\left(\frac{4276}{x_{17}}-3.961\right)\frac{x_{23}Rx_{17}}{dV_a}\frac{x_{21}Rx_{17}}{dV_a}-\frac{x_{24}Rx_{17}}{dV_a}\frac{x_{22}Rx_{17}}{dV_a}\right]-x_6$$

$$\dot{x}_{22}=A_a\frac{u_2}{Ru_3}u_4u_7-A_ax_{27}\frac{x_{22}}{dV_a}+3AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)$$

$$+AK_s\left[\exp\left(\frac{4276}{x_{17}}-3.961\right)\frac{x_{23}Rx_{17}}{dV_a}\frac{x_{21}Rx_{17}}{dV_a}-\frac{x_{24}Rx_{17}}{dV_a}\frac{x_{22}Rx_{17}}{dV_a}\right]-x_2$$

$$\dot{x}_{23}=A_a\frac{u_2}{Ru_3}u_4u_8-A_ax_{27}\frac{x_{23}}{dV_a}+AK_r\left(\frac{x_{20}Rx_{17}}{dV_a}\right)^{\alpha}\left(\frac{x_{21}Rx_{17}}{dV_a}\right)^{\beta}\exp\left(-\frac{E_r}{Rx_{17}}\right)$$

$$-AK_s\left[\exp\left(\frac{4276}{x_{17}}-3.961\right)\frac{x_{23}Rx_{17}}{dV_a}\frac{x_{21}Rx_{17}}{dV_a}-\frac{x_{24}Rx_{17}}{dV_a}\frac{x_{22}Rx_{17}}{dV_a}\right]$$

$$\dot{x}_{24}=A_a\frac{u_2}{Ru_3}u_4u_9-A_ax_{27}\frac{x_{24}}{dV_a}+AK_s\left[\exp\left(\frac{4276}{x_{17}}-3.961\right)\frac{x_{23}Rx_{17}}{dV_a}\frac{x_{21}Rx_{17}}{dV_a}-\frac{x_{24}Rx_{17}}{dV_a}\frac{x_{22}Rx_{17}}{dV_a}\right]$$

$$\dot{x}_{25}=A_c\frac{u_{10}}{Ru_{11}}u_{12}(1-u_{13})-A_cx_{28}\frac{x_{25}}{dV_c}$$

$$\dot{x}_{26}=A_c\frac{u_{10}}{Ru_{11}}u_{12}u_{13}-A_cx_{28}\frac{x_{26}}{dV_c}-x_4$$

$$\dot{x}_{27}=\frac{1}{\sum_{20}^{24}(x_iM_i)}\left[A_a\frac{u_2}{Ru_3}\sum_5^9(u_iM_i)u_4^2-A_a\frac{\sum_{20}^{24}(x_iM_i)}{dV_a}x_{27}^2-\right.$$

$$\left.x_2M_{H_2}x_{27}+A_a\Delta P_a-A\tau_{w,a}-x_{27}\sum_{20}^{24}(\dot{x}_iM_i)\right]$$

$$\dot{x}_{28}=\frac{1}{\sum_{25}^{26}(x_jM_j)}\left[A_c\frac{u_{10}}{Ru_{11}}((1-u_{13})M_{N_2}+u_{13}M_{O_2})u_{12}^2-A_c\frac{\sum_{25}^{26}(x_jM_j)}{dV_c}x_{28}^2\right.$$

$$\left.-x_4M_{O_2}x_{28}+A_c\Delta P_c-(A_1+A_2)\tau_{w,c}-x_{28}\sum_{25}^{26}(\dot{x}_iM_i)\right]$$

输出：

$$y_1=\frac{u_1}{u_1+R_o}x_1$$

$$y_2=\frac{x_1}{u_1+R_o}$$

$$y_3=\frac{\sum_{20}^{24}x_iRx_{17}}{dV_a}$$

$$y_4=x_{17}$$

$$y_5=x_{27}$$

$$y_6=\frac{x_{20}}{\sum_{20}^{24}x_i}$$

$$y_7=\frac{x_{22}}{\sum_{20}^{24}x_i}$$

$$y_8=\frac{x_{22}}{\sum_{20}^{24}x_i}$$

$$y_9=\frac{x_{23}}{\sum_{20}^{24}x_i}$$

$$y_{10}=\frac{x_{24}}{\sum_{20}^{24}x_i}$$

$$y_{11}=\frac{\sum_{26}^{25}x_jRx_{18}}{dV_c}$$

$$y_{12}=x_{18}$$

$$y_{13}=x_{28}$$

$$y_{14}=\frac{x_{26}}{\sum_{26}^{25}x_j}$$

$$y_{15}=u_{14}\frac{x_{19}}{u_{15}}$$

$$y_{16}=x_{19}$$

$$y_{17}=u_{16}$$

用一个紧凑的结构,状态空间模型可以表示为

$$\begin{aligned}\dot{\boldsymbol{x}}&=f(\boldsymbol{x},\boldsymbol{u})\\ \boldsymbol{y}&=g(\boldsymbol{x},\boldsymbol{u})\end{aligned}\tag{7.80}$$

7.7.3 模型验证

建模过程的重要步骤之一是验证模型。因为文献中没有完整的动态实验数据，将通过两个步骤来验证模型：第一步是通过 Wagner 等人（1998）以及 Tsai 和 Barnett（1997）提供的静态 *V-I* 实验结果的比较验证转换和扩散过程；第二步是用已出版的其他研究者的仿真结果比较电压的模拟动态响应。

Wagneret 等人（1998）研究了扩散的影响，以及通过 EIS 博德图观察到的动态实验属性。在湿度为 3% 的 H_2、纯 O_2 和无限大负载电阻 R_{load} 下，在 1223K 的温度下，模拟稳态开路电压（OCV）是 1117mV。这与 Wagner 等人（1998）给出的相同条件下的测量结果 1114mV 很吻合。因为 Wagner 等人（1998）没有透露扩散层的厚度，所以他们实验所得到的 *V-I* 图不能用于验证模型。

将用开发的模型模拟稳态 *V-I* 输出与 Tsai 和 Barnett（1997）给出的实验结果进行比较，如图 7.11 所示。

尽管 Tsai 和 Barnett（1997）没有直接提供扩散层的厚度，但可以间接推断出这些信息。当 I 到达极限电流密度 I_1，天然气供应率最大，扩散层的厚度可以由下式计算：

$$I_1=4FJ_{O_2}^r=4FAD_{O_2}\frac{C^b-0}{l_c}\tag{7.81}$$

根据式（7.81），计算扩散层的厚度 $l_c=1.26\times10^{-3}\mathrm{m}$。

在阳极侧，扩散不是整个运输过程的限制因素，因为阳极的厚度也没有提供，简单地假设 l_a 与 l_c 相同。

内阻可以由 Tsai 和 Barnett（1997）的 EIS 图直接读取。弧和实轴之间的第一个交点是固有阻力。从 EIS 图可以看出，在 1123K 下 $R_o=0.16\Omega$，$R_{ct}=0.16\Omega$。

V-I 图显示，与预期的一样，模拟结果总体与实验数据一致，有些许差异。造成差异的主要原因是用于模拟的模型参数误差，或者是存在没有考虑进模拟的某些过程。这些数据来源不同，大范围地分布于不同的文献。

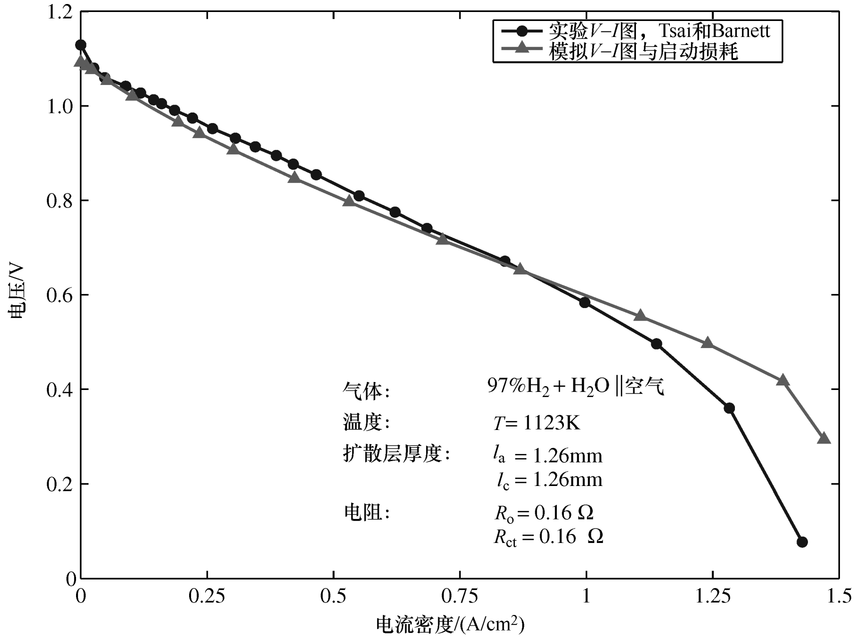


图 7.11　模拟和实验 V-I 图对比

图 7.12 比较了不同研究人员开发的不同模型中，模拟电压瞬变响应负载的阶跃响应试验。这里比较的只是那些考虑了电池温度的模型。在本章中可以看到，模型的整体动态趋势与其他人的是一致的。在这里的模型中，不仅考虑质量/热/动量

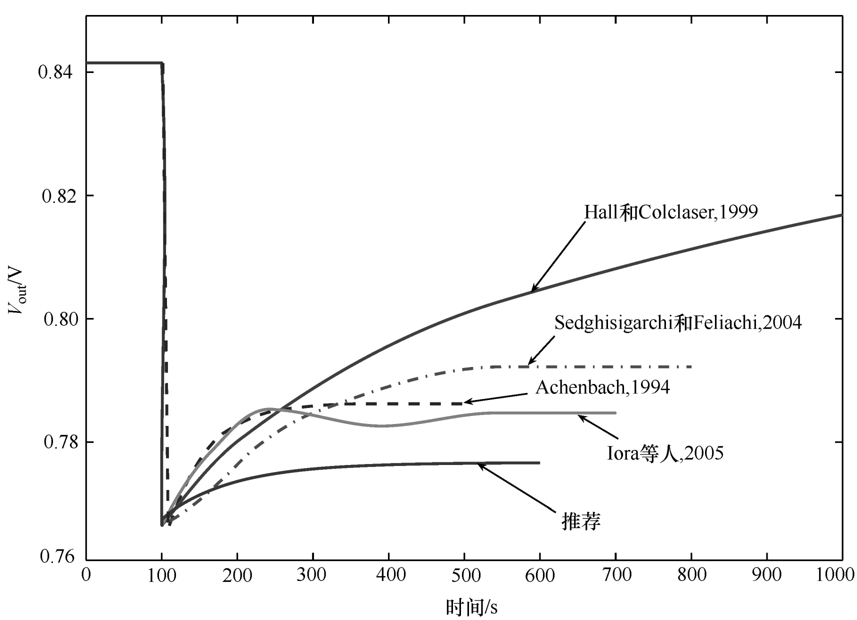


图 7.12　负载阶跃电压响应比较

转移、内部转换和转移反应，还考虑了转换反应所吸收的热量。因此，模型中的模拟稳态电压比别人的要低，电池温度也要低。

7.8　模拟动态属性研究

有些性质很难用物理实验来研究，可以用这个模型来研究。在下面的几节中，模拟了输入变化对 SOFC 的影响。

阶跃响应测试广泛应用于过程动态的研究，它们可以揭示关键过程动态参数，如时间常数、增益和时间延迟。仿真结果显示了各种变量对电池的性能有什么样的影响，包括动态的和静态的。从仿真获得的知识在 SOFC 的过程和控制设计中是很有用的。

7.8.1　扩散动态

为了研究过程的动态性质，可以选择模拟条件来揭示感兴趣的特定的动态。在本节中，研究扩散过程和扩散层的影响。

在气体中，分压是一个干扰，会导致输出电压的波动，同时也影响反应速率。图 7.13 显示了由于氢气压力变化导致的阶跃响应。仿真结果表明，气体膨胀带来的压力变化对电压和电流影响相对较小。

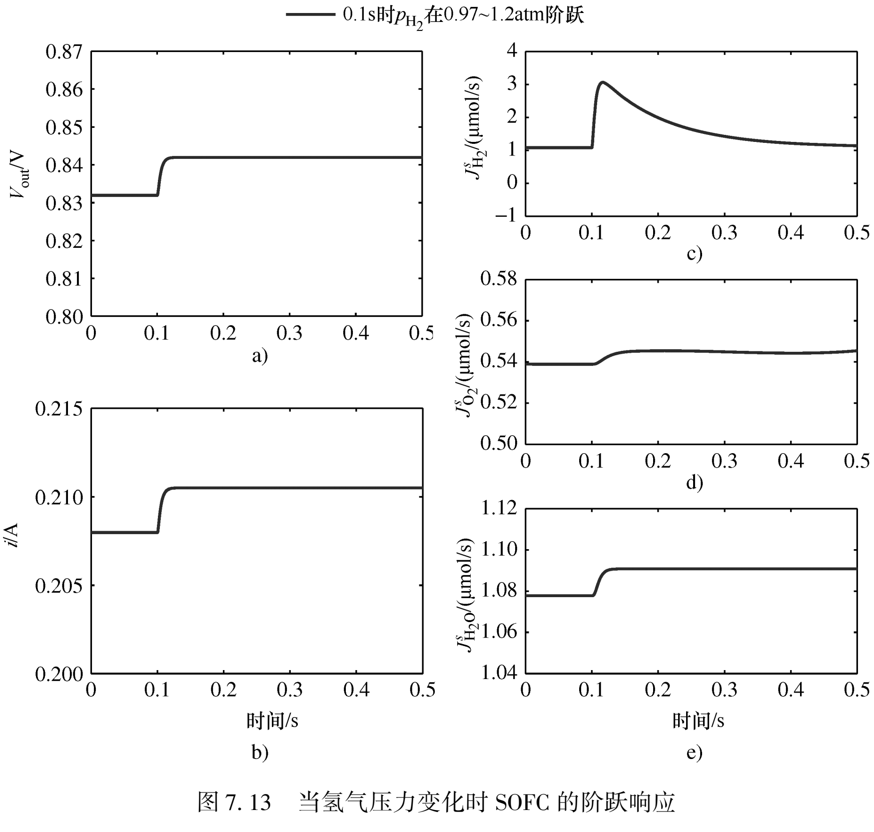


图 7.13　当氢气压力变化时 SOFC 的阶跃响应

当负载电阻下降到一个更小的值时，将从 SOFC 中带走更多的电流。因此，在三相反应界面反应物的消耗速率将会大大提高。由于扩散层的存在，三相反应界面处，反应物的分压显著减少，产生更高的浓度梯度，这样会有更多的反应物扩散到三相反应界面。如此低的分压会影响电压，以至于电池在主要浓度损失区运行。*V-I*图中，在高电流区域主要浓度损失区会下降，电压明显低于直线电阻损失 *V-I* 区域。由图 7.14 可见，主要电阻损失区域和主要浓度损失区的动态也有明显的不同，似乎是高阶动态。

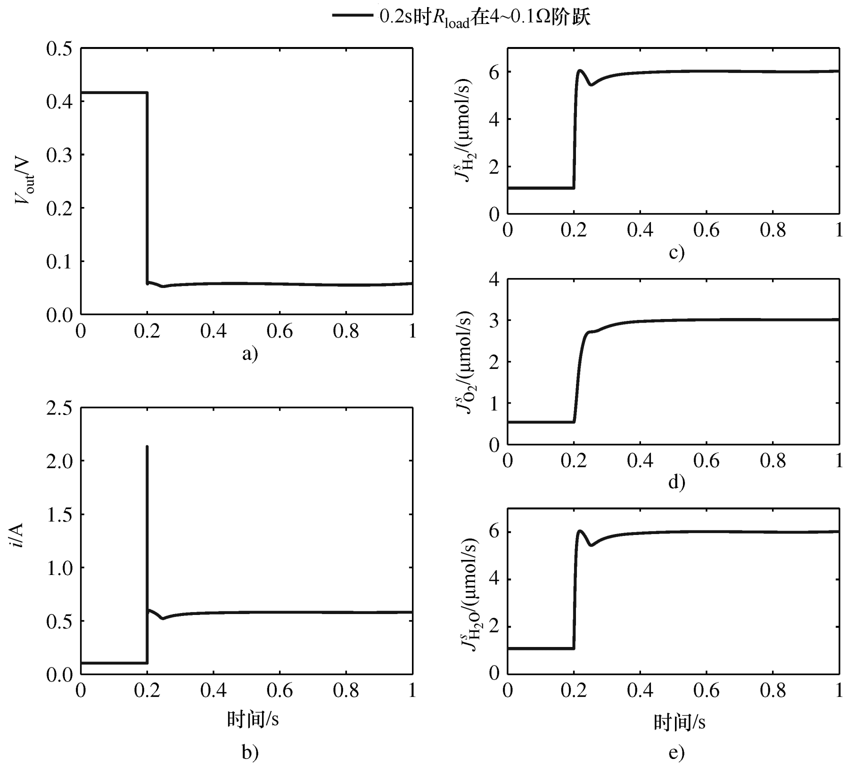


图 7.14　当进入浓度损失区域时 SOFC 的阶跃响应

通过仿真，可以发现浓度损失主要源于氧气供应过程，由于氧的扩散系数比氢和水要小，所以氧气扩散在 SOFC 的动态过程中起主要作用。

扩散层的厚度对 SOFC 的动态属性有强烈影响。随着扩散层厚度的增加，动态响应的时间常数也会增加，扩散阻力也增加。图 7.15 显示了影响。燃料电池的扩散层不仅包括电极层还包括边界层。边界层的厚度受到流体状态，尤其是流速的影响。所以扩散层的厚度可能相差很大。

总的来说，通过仿真可以得到这样的结论，扩散对燃料电池的静态行为有影响，但对动态行为影响还是较小的。扩散过程导致的时间常数大约只有几百毫秒，

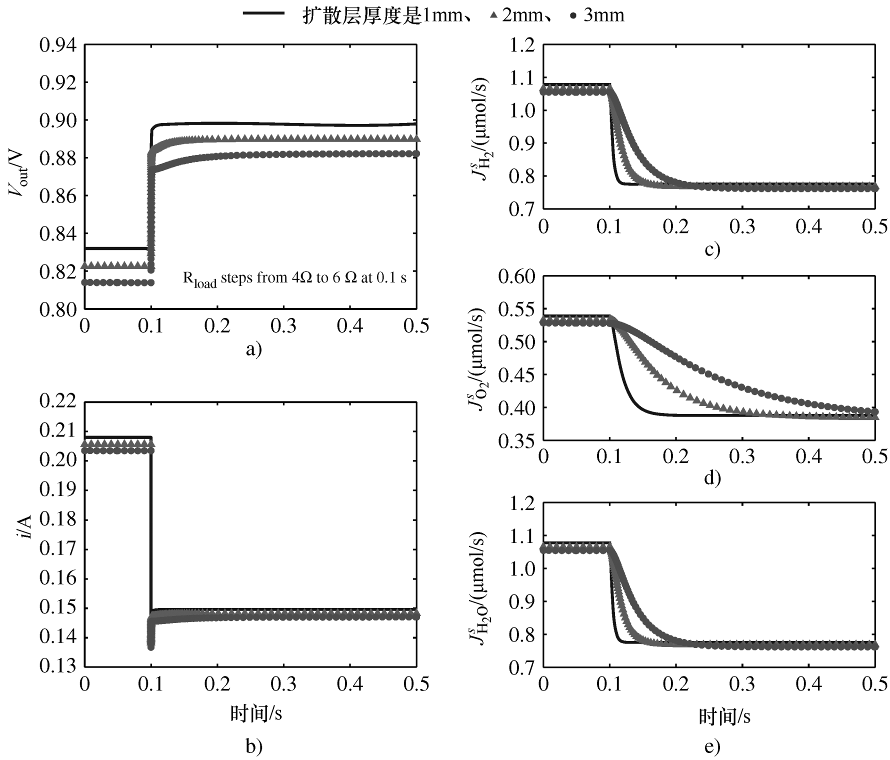


图 7. 15　SOFC 阶跃响应的扩散层厚度效应

而静态效应主导着浓度损失。

7. 8. 2　燃料输送过程动态

燃料输送对动态响应的影响是分析中的主要焦点之一。对于进气口处 P_{fuel}、T_{fuel}和 u_{fuel}大约 5% 的变化的阶跃响应，在图 7. 16 和图 7. 17 中做了仿真和比较。

随着燃料入口处的压力 P_{fuel}的增加，CH_4 和 H_2O 在入口处的浓度也立即增加。因此，转换反应会更快，从电池吸收更多的热量，从而逐渐降低了电池的温度，如图 7. 16c 所示。燃料进气口压力越大，意味着 H_2 的浓度越大，同时，转换反应越快，产生 H_2 速率也越快。较高的 H_2 浓度会导致输出电压增加。因为电池温度逐渐降低，电压也会略有下降，如图 7. 16a 所示。因为外部负载没有变化，电流随着输出电压而变化，如图 7. 16b 所示。

当入口燃料的温度增加时，放热转换反应的平衡点向相反方向 CO 和 H_2O 移动。H_2 和 CO_2 的浓度会减小，这就导致输出电压和电流降低，如图 7. 16a 和图 7. 16b所示。内部电阻消耗的电能会更少，产生的热量也会更少。转换反应的平衡点的变化也导致了总压强的轻微变化。另一方面，很高的燃料温度会加快转换反应的进行，反应从电池吸收热量也会更快，因此电池温度就会

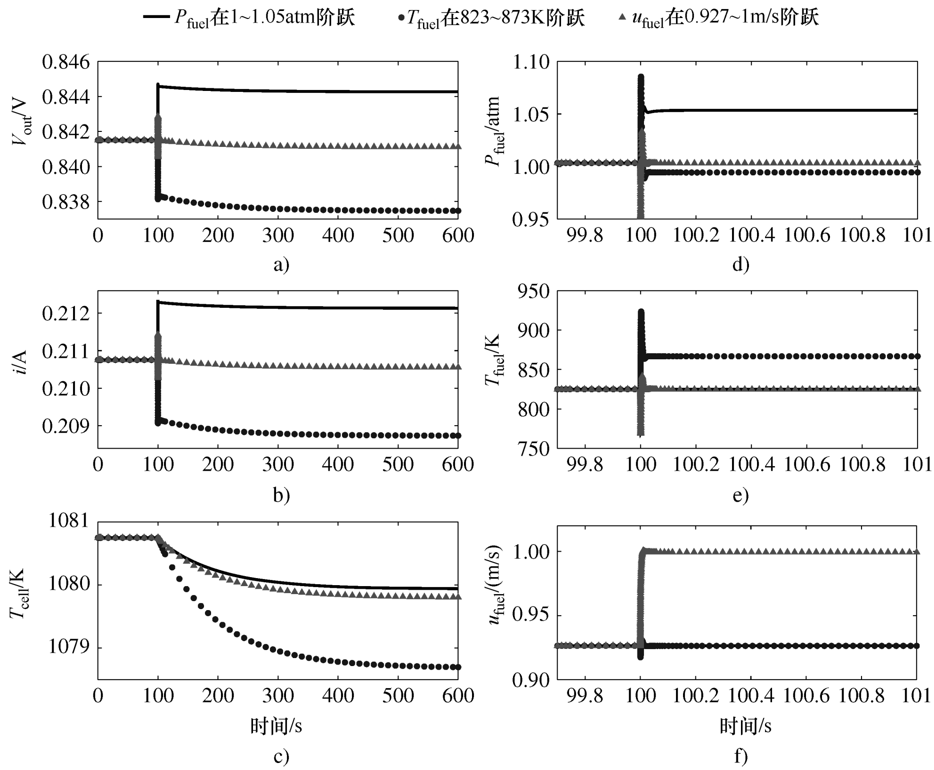


图 7.16　增加入口燃料流的压力、温度和速度时 SOFC 的阶跃响应

降低。

燃料流速 u_{fuel}的改变也会影响 SOFC 的性能。燃料流速越快，传热系数 h_a 越大，因此会有更多的流量从电池转移到燃料流。燃料温度就会增加，而电池温度就会降低。电池温度的下降会略微影响输出电压和电流。u_{fuel}对其他变量的影响相对较小。

由于电池具有相对大的热容，电池温度的反应要比其他变量的反应慢得多。电池温度的时间常数大约在几百秒左右，而其他变量大约在不到 1/10s。电池温度是一个重要的变量，它几乎影响到电池动态的每一个方面，它较慢的动态特性出现在对其他变量的反应中。从控制应用程序的角度看，由于快和慢之间区别很大，更快的动态可以视为一个即时的变化，在动态分析中只需要考虑慢动态。

一般来说，入口压力 P_{fuel}和燃料温度 T_{fuel}对输出功率有很大的影响。流速 u_{fuel}扰动的影响相对较小。入口燃料温度的增加会降低输出电压，会出现反方向的转移反应。这就导致氢分压的减小，从而减小输出电压。此外，入口燃料温度的增加也加快了转换反应，反过来这会吸收更多的热量，导致温度和电压的降低。因此，为

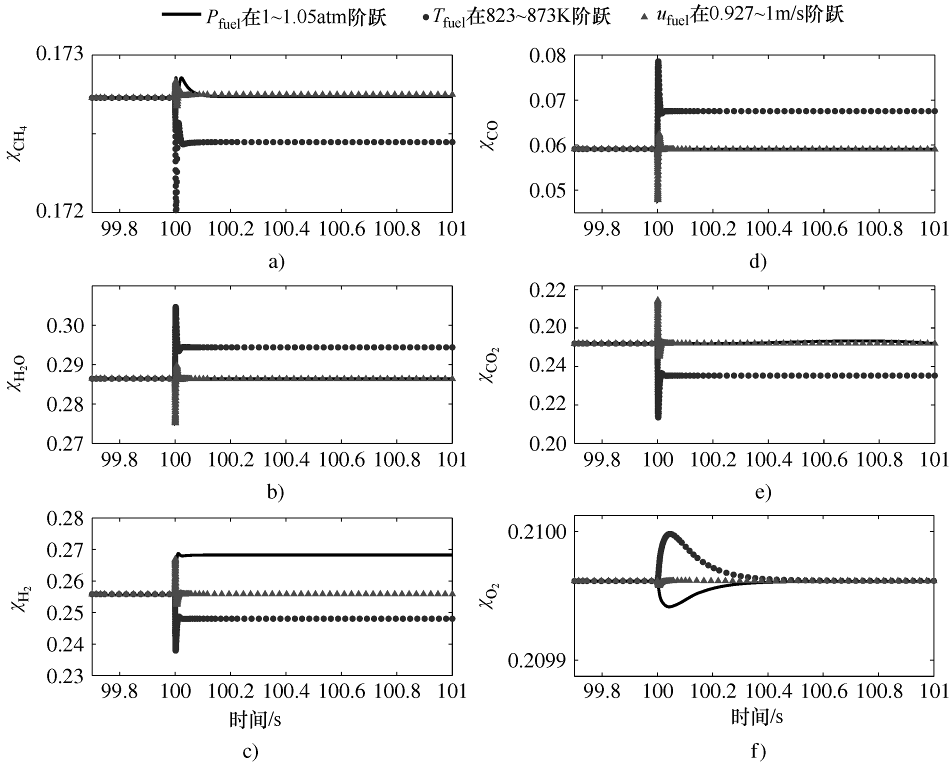


图 7.17　增加入口燃料流的压力、温度和速度时摩尔分数的阶跃响应

了达到最佳性能，优化 T_{fuel}很重要。

7.8.3　空气输送过程动态

通过 P_{air}、T_{air}和 u_{air}逐级测试，研究空气运行条件对燃料电池性能的影响。

当阴极侧气流入口压力 P_{air}从 1atm 增加到 1.05atm 的时候，阴极侧流动管道内 CO_2 的浓度相应地增长，三相反应界面 CO_2 的浓度也会增加。反应物浓度增加，输出电压 V_{out}和电流 I 就会增加。由于空气中氧气的摩尔分数大约为 21%，所以输出电压和电流的增加并不明显，如图 7.18a 和图 7.18b 所示。因此固有阻力产生的热量较少，对阳极侧燃料特性几乎没有影响。

为了加热电池，通过设计，电池中空气的温度比电池温度高。入口处空气的温度直接影响电池的温度。空气流与电池的温差越大，就会有越多的热量转移给电池，最终电池温度会增加到一个很大的值，如图 7.18c 所示。电池温度的增加会导致阳极侧一系列状态的改变。会有更多的热量从电池传递到燃料，从而增加了燃料的温度 T_{fuel}和压力 P_{fuel}。此外，转移反应的平衡点向反方向移动，因此 H_2 和 CO_2 的浓度就会降低。但是输出电压仍将增大，因为电池温度对电压的影响比转变反应变化的影响大。可以看到空气流的温度对动态响应有很大的影响，如图 7.18 和

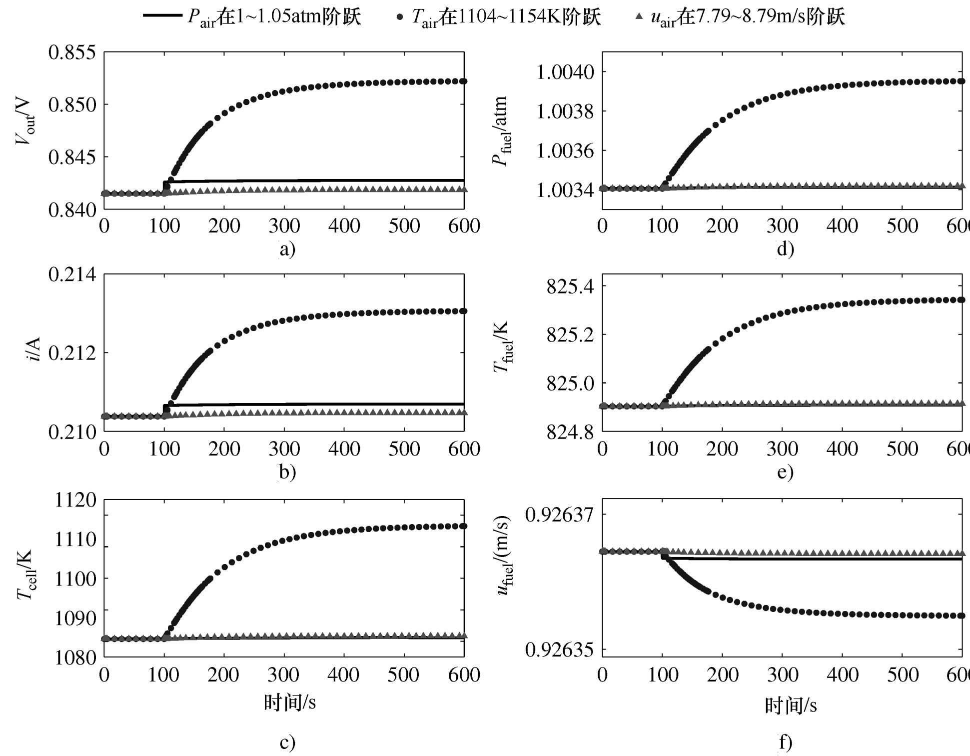


图 7.18　增加进气空气流的压力、温度和速度时 SOFC 的阶跃响应

图 7.19所示。

空气流速的阶跃变化主要影响热传递系数 h_c，从空气传到电池的热量也会增加。因此电池温度略有增加，但是对其他变量几乎没有影响。

与其他操作条件相比，入口处空气的温度对阴极侧固相温度有很大的影响，这样就会影响电池的性能，包括静态的和动态的。

7.8.4　外部负载动态

在正常操作条件下，外部负载是一个外部变量，与电池没有关系。因为它的波动对电池的动态有最直接和最重大的影响，所以需要研究其影响。

图 7.20 和图 7.21 描述了 V_{load}、I、T_{cell}、P_{fuel}、T_{fuel}、u_{fuel} 的阶跃响应和物质的摩尔分数对外部负载 R_{load}的阶跃变化。

当内阻从 4Ω 降低到 2Ω 时，电流 I 立即增大，随后略有下降，然后约 0.2s 的时间常数会有一个缓慢变化，如图 7.20b 所示。电流的激增来源于存储在双层电容器中的电子，激增之后相对缓慢的反应主要源于扩散过程，这个已在前面内容中讲过。

前面讨论的两个快速响应，所有过程变量中，还有一个响应要慢得多，大约为 200s。这么慢的响应主要是由于温度动态反应缓慢。当负载 R_{load}减小时，

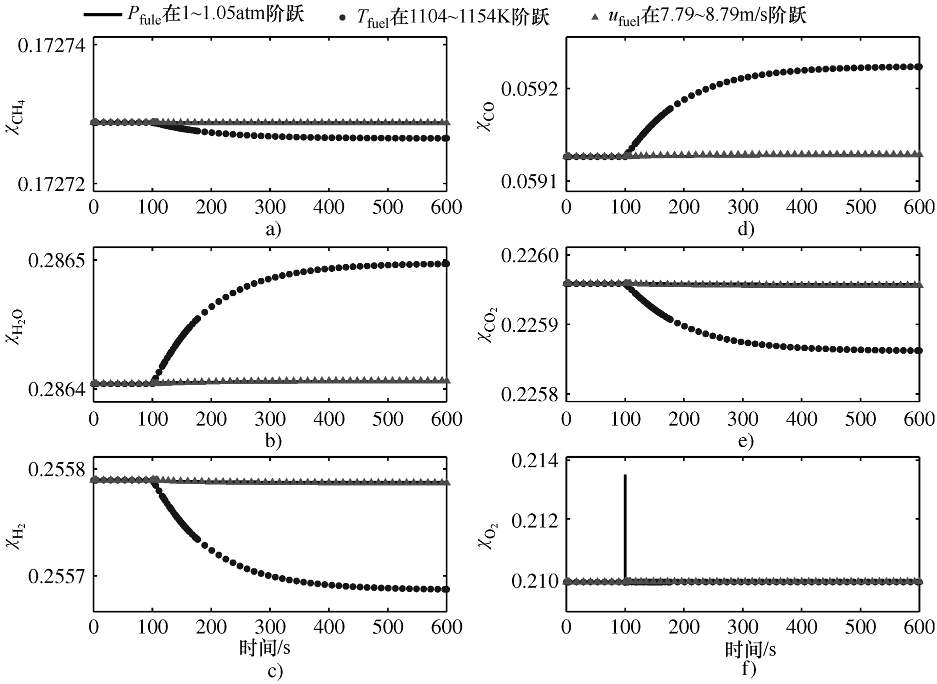


图 7.19　当增加入口空气流的压力、温度和速度时的摩尔分数的阶跃响应

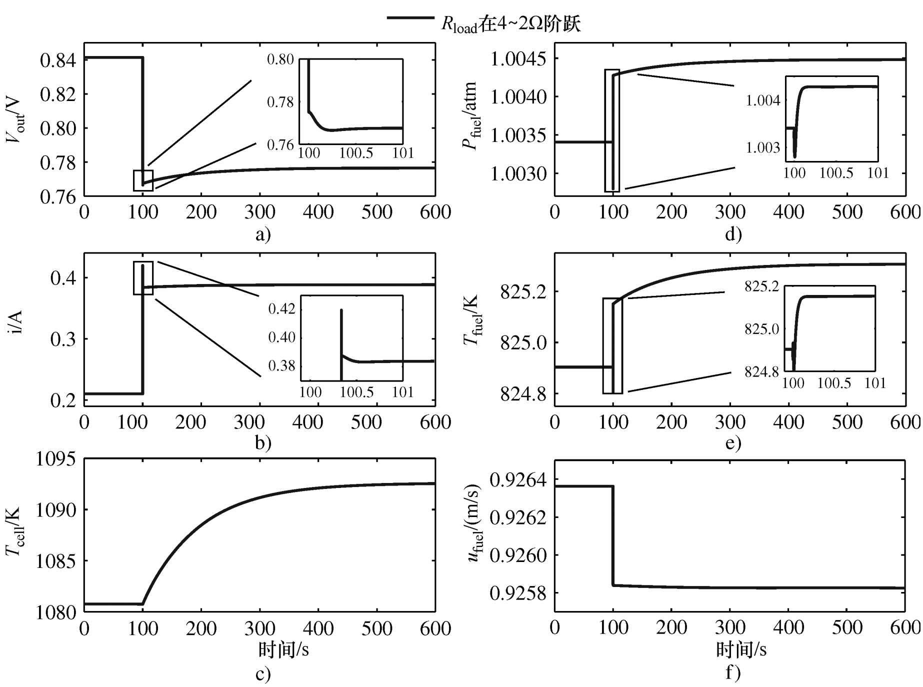


图 7.20　当 R_{load} 改变时 SOFC 片的阶跃响应

电流会增加，内阻消耗的电能也增加。这部分能量转化为热量，直接让电池升温。然而，由于固态电池材料有很大的热容，电池响应相对较慢，如图 7.20c 所示。

如图 7.20d ~ f 所示，负载的阶跃变化对燃料流量也有影响。当需要的电流增大时，就需要有更多的反应物 $J^s_{H_2}$ 和 $J^s_{H_2O}$ 流入通道。这就改变了燃料流中每种物质的摩尔分数，反过来影响重整和转换反应。放热的转移反应会向反方向进行，这就导致燃料温度快速下降。然而，当热的 H_2O 进入燃料通道时，温度又会上升，如图 7.20e 所示。温度和 H_2O 浓度的变化会带来总压强和 CH_4 浓度的变化，分别如图 7.20d 和图 7.21a 所示。其他物质的摩尔分数的响应如图 7.21b ~ e 所示。

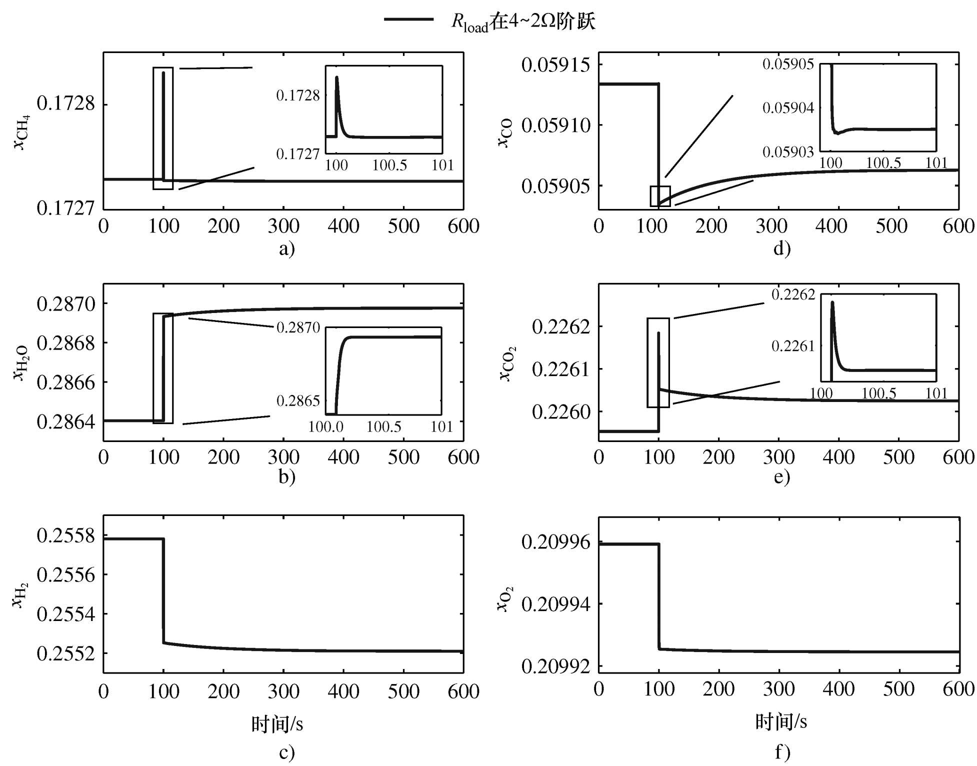


图 7.21　当 R_{load} 改变时物质摩尔分数的阶跃响应

当负载减小时，会需要更大的电流，燃料流中 H_2O 会取代 H_2，这导致燃料流量的密度的增加。根据动量平衡，流速会增加，如图 7.20f 所示。

仿真结果表明，外部负载的波动对 SOFC 的快速和慢速动态响应都有很大的影响。较慢的动态可以通过适当的控制措施来调整，但是快速动态可能必须通过适当的外源设备或过程的设计来调整。

7.9 注解和参考文献

如果读者想进一步了解 SOFC，推荐阅读 Achenbach（1994，1995）、Padullés 等人（2000b）、Zhu 和 Tomsovic（2001）、Sedghisigarchi 和 Feliachi（2004a）、Xue 等人（2005）、Gemmen 和 Johnson（2005）以及 Iora 等人（2005）的文献。这些方法通常是从集总参数模型开始，通过能量守恒定律，然后通过流动过程引起的传质动态，最后是有内部反应引起的各物质动态。本章总共考虑了流体动态、能量和质量转移、扩散和内阻和重整/转移反应。本章主要参考了 Qi 等人（2005） and Qi 等人（2006）的文献。

生成焓的更详细的知识可以在“Thermodynamics”（Smith 和 Ness，1987b）一书中查到。热传递的知识如传导、对流、辐射可以在“Heat Transfer”（Welty 等人，1984）一书中查到。改革和改变反应的主要参考文献是 Ahmed 和 Foger（2000）以及 Campanari 和 Iora（2004）的文献。有关流体压力梯度的知识可以在 Blackburn 等人（1960）的文献中找到。流体通道的摩擦系数以及视图因素是根据“Chemical Engineering Handbook”（Perry 等人，1997a）确定的。

模拟中所用的输入条件和力学参数是参考了 Campanari 和 Iora（2004）的文献；这些参数都在附件表 A-1 和表 A-2 中做了总结。固相材料的特性参数主要是基于 Hall（1997）和 Bessette（1994）的文献。改革和转变反应的反应动态参数是基于 Ahmed 和 Foger（2000）的文献。所有的这些参数在表 A-2 中做了总结。

CH_4、H_2O、H_2、CO、CO_2、O_2、N_2 和空气的性质与温度有很大的关系，这些关系不能忽视。在附录中，这些气体性质用温度的多项式表示了一下。属性参数和温度的相关性参考了“Chemical Engineering Handbook”（Perry 等人，1997a）一书。

第 8 章　管式 SOFC 动态模型：简化的第一性原理方法

8.1　简介

与用于仿真和分析的模型不一样，用于控制的模型不仅需要接近于系统的整体动态，还要有明确的输入/输出的关系。与此同时，令人满意的是该系统的要求相对较低，便于控制的设计。在过去出现了几种通过减少模型的技术来简化复杂动态模型的方法，可是利用单纯的模型简化方法很有可能漏掉重要的物理信息，所以从第一性原理中直接提出一种更简化的模型可以更好地解决这个问题。

在前几章化学工程第一性原理的基础上，本章将要介绍一个关于管式固体氧化物燃料电池（SOFC）堆栈的低阶模型。像气相混合、能量转换的快速过程被称为稳态模型，像固相对温度的缓慢过程被称作动态模型。因此，导出的这个用来解释 SOFC 中物理过程的低阶动态模型，并不是简单地从前面章节中通过模型减少技术而建立的高阶模型中简化的，而是通过考虑所有机制直接利用第一性原理得到的模型。

本章将阐述如何在第一性原理基础上发展建立一个简化或集总参数的动态模型。通过模拟将简化的模型与第 7 章中的复杂模型进行比较。本章简单回顾一下指导简化模型发展的 SOFC 原理。

8.1.1　过程变量的关系

从燃料和空气的输入到电的输出，燃料电池中的能量转换是一个复杂的化学和电化学过程。正如第 7 章中阐述的质量、能量和动量的转换过程和化学反应都同时发生。这些过程中，包括它们的物理特性，例如，压力、温度和它们的相互影响关系。在第 7 章中介绍的 SOFC 原理的基础上，过程变量关系的 SOFC 原理总结在图 8.1中，它可以用作将来发展简化第一性原理模型的一个指导原则。

在燃料和空气流动通道中，入口和出口的压力决定着流速。这些流被称为主体流，它们影响着堆栈中的热交换过程，例如电池的温度，这些在图 8.1 中有描述。在另一方面，SOFC 产生的电流由外部负载决定，因此也通过电化学反应决定着反应物的消耗速率。总压和各物质的分压由主体流的速率和反应物的消耗速率共同决定。SOFC 能够输出的电压受反应物的分压和温度影响，因此反过来看，同样影响着输出的电流。

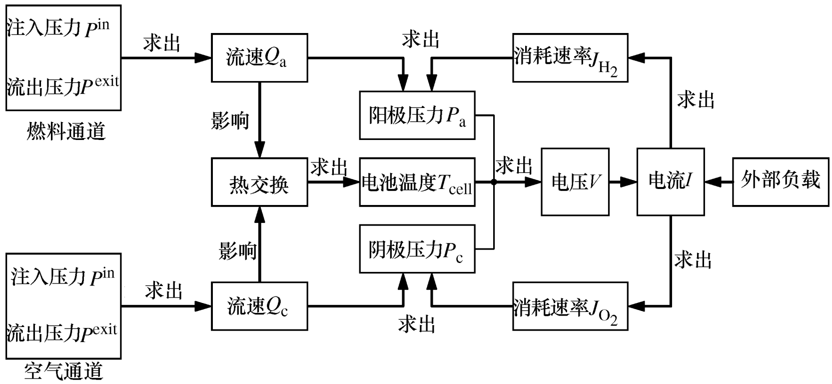


图 8.1　SOFC 的原理逻辑关联流程变量

8.1.2　功率输出限制

正如第 7 章讨论的，一个 SOFC 的功率受多个限制影响。第一个限制因素就是反应物供应率。在液流通道里的反应物必须穿过多孔电极到达反应点。这个过程受 tpb 和主要气流体积的浓度梯度控制。在电流需求增加的时候，在 tpb 中的氢气和氧气的浓度减少而形成更大的浓度梯度。一旦其中一个减少到零，则反应物供应率达到最大值。此时，电流不能进一步增加，这就是燃料电池能够提供的最大电流值。在这种条件下，正如 Nernst 方程描述的，电压降低至零。

第二，当反应物供应率足够时，反应速率和反应中得到的可用的反应点将会限制电流的产生速率。在大多数例子中，在阳极的反应是相当快的。但是，在阴极上产生离子的反应相对较慢（Mitterdorfer 和 Gauckler，1999），所以输出电流也受最大产出离子速率的限制。

第三，最常见的情况就是当燃料电池在工作时反应物供应率和反应速率都充足，输出电流受外部负载的阻抗限制。在这个例子中，电流是决定电化学反应和扩散的共同因素。

第四，电流也受电解液的离子电导率和本征电阻的限制。它们服从欧姆定律，而且通常被共同看作内阻。当电压低于内阻跟燃料电池产生的电位相等时，输出电压减少到零。在常规的操作环境下，反应不能到达之前提到的限制条件。因此，电流主要受欧姆定律的控制。

8.2　SOFC 堆栈的低阶状态空间模型

跟其他许多过程一样，一个 SOFC 堆栈的操作应该考虑几个安全因素：第一，当在一个相对较高的温度下操作时，在阳极和阴极之间的温度和分压差别必须维持在能够保证其完整结构的设计范围内；第二，燃料电池的输出电压为了持续使用必

须稳定；第三，燃料电池在使用时必须达到一个最佳的效益成本比。

8.2.1　物理过程

SOFC 的物理过程总结成一个结构图，如图 8.2 所示。

燃料引入和空气流动的特性用几个状态描述，例如压力、温度、速率、焓变和组成。通过在通道的界面条件，主要的质量流动速率由堆栈的进气性能和几何形状决定。二次流沿着流动通道发生是因为从主流体流向的反应位点的反应物消耗减少。如果用合成气作燃料，重整和转移反应一样能够在阳极流道发生。这些反应改变了燃料气体的组成。所有这些因素共同决定着所有燃料的压力和在流通通道里各组分的分压。这个过程在图 8.2 中的堆栈和通道部分。

主要流体的分压和输出电流决定着扩散过程，同时决定反应位点邻近的分压。它们之间的关系表示在图 8.2 中的电池部分。扩散过程形成了二次流。反应位点邻近的分压决定着电动势（EMF）和电化学反应速率。

能量也通过流道进行转移，同时伴随着质量流和化学反应。由质量流所携带的主要能量形式是焓。反应物扩散进入电池参与电化学反应，然后释放能量。吉布斯自由能转化成电能，剩余的生成焓以反应热的形式释放。该电能的一部分被供给到外部负载，另一部分由本征电阻消耗，并且被转换为热量。

由于温差，使电池用燃料和空气进行热交换，并由电池中输送热到燃料和空气流。燃料和空气流的温度影响传热过程和电池的温度，电池也给转化反应提供热。能量流和转换过程都显示在图 8.2 中，在堆栈/通道区段。

8.2.2　建模假设

要开发一个简化的低阶 SOFC 堆栈模型，必须通过进行以下假设忽略一些快速的或微不足道的物理过程：

1）供给到 SOFC 堆栈的燃料是预重整甲烷和空气；
2）SOFC 堆栈的出口是开放的，因此，出口压力为 1atm；
3）燃料的温度、压力和流速在流动通道中是均匀的；
4）堆栈的所有单元具有相同的温度；
5）每一个电池单体的电压是相同的；
6）电池在堆栈以串联的方式连接；
7）CO 不会参与在电化学反应中；
8）堆栈和周围环境之间没有热传递发生。

8.2.3　I/O 变量

本节的主要目的是开发一种用于控制应用的 SOFC 堆栈简化动态模型。因此，只有确切的、可测量的或可操作的物理变量被指定为 I/O 变量。

8.2.3.1　输入变量

从物理的角度看，SOFC 堆栈的输入是燃料/空气流。入口流量的属性由它们的压力、温度、流率和组成共同决定着。

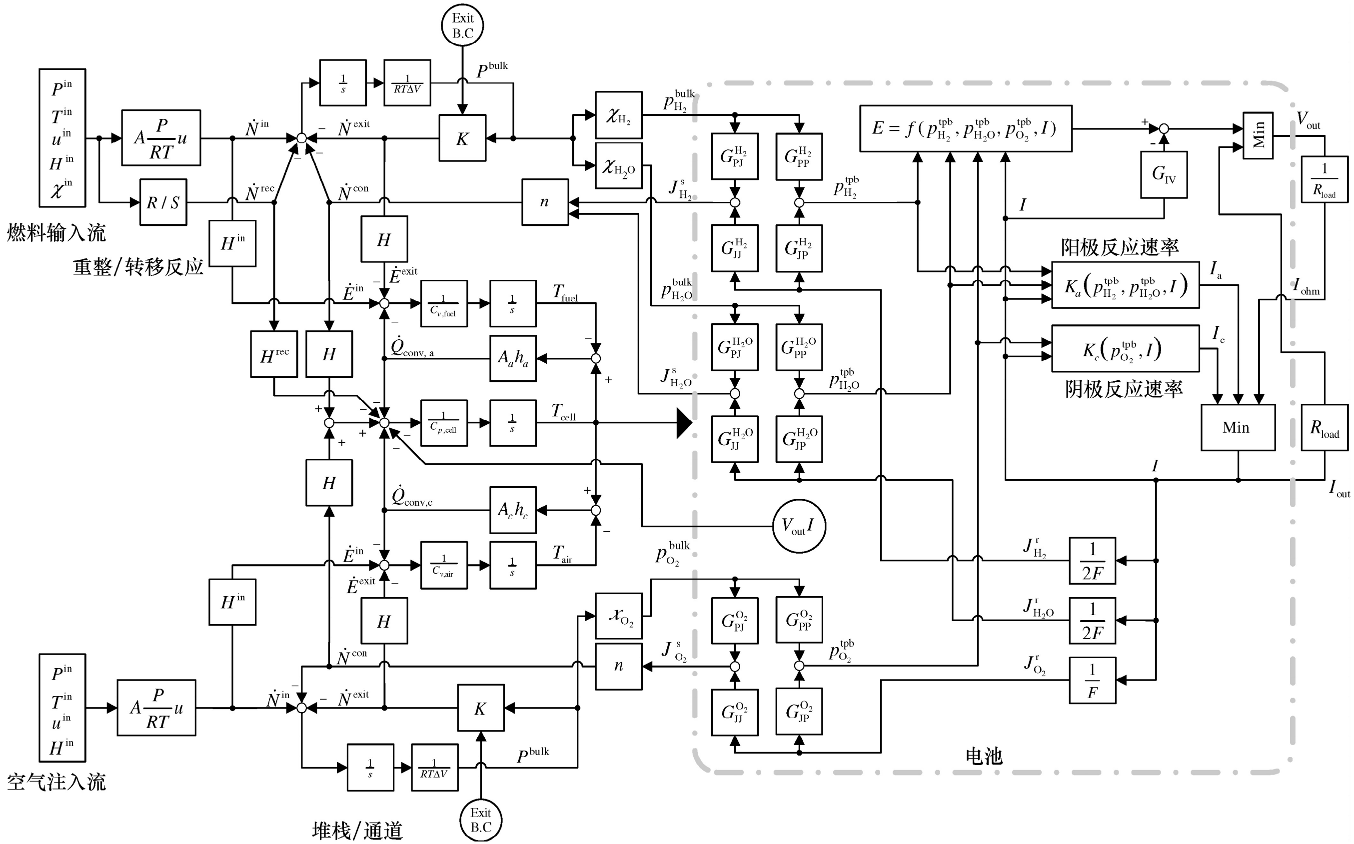


图 8.2　燃料电池的物理过程框图

SOFC 堆栈的输出是电，它是由电压和电流决定的。电压与燃料的压力有关，电流与燃料流率有关。为了满足当前的负载需求，需要调节燃料和空气流的压力和速率。

因为假设堆栈的出口是打开的，所以出口处的压力是固定的。因此，堆栈的压力和燃料/空气流的流速都是由入口的压力、P_{fuel}^{in}和 P_{air}^{in}、流阻共同决定的。因此P_{fuel}^{in}和 P_{air}^{in}被作为输入变量的选择。

由 Padullés 等人（2000b）提出的，SOFC 堆栈的出口也可以用来调节电池堆栈的压力和流速。但是这种方法有两个缺点：第一，废气的压力会有波动，这是不适合输送到涡轮发动机进一步利用的；第二，这将需要更高的进气流的压力，以保证足够的工作范围。

8.2.3.2　输出变量

燃料电池的输出电压 V_{out}是一个必须控制在一定范围内并且需要稳定的重要变量。它被认为是输出变量中的一个。

由于内阻，电池温度 T_{cell}比周围的温度高。它可以显著影响 SOFC 堆栈的性能。T_{cell}过高会引发堆栈的故障。因此，T_{cell}也必须密切监控，并加以控制。这是本章中考虑的第 2 个输出变量。

8.2.3.3　扰动变量

电流 I 是由电压和负载阻抗确定的，它反过来又决定了电化学反应速率和堆栈的总体性能。电能分配网络中的负载阻抗可能不被控制，可以看作一个干扰。然而，由于电压 V，电流和负载阻抗都不是独立的变量，所以任何一个都可以被视为一个扰动变量。如果负载阻抗被认为是一个扰动变量，电流 I 可以被视为一个输出量，跟在第 7 章处理的那样。但是，在本章中，I 被看作一个扰动变量。

入口温度的波动，即 T_{fuel}^{in}和 T_{air}^{in}影响堆栈的性能表现。它们不能被控制，因此，它们被指定为扰动变量。其余的变量，燃料成分χ^{in}，被视为系统参数。

I/O 变量被总结在表 8.1 中。

表 8.1　模型 I/O 输出变量

输入	输出	扰动
P_{fuel}^{in}（atm）	V（V）	I（A）
P_{air}^{in}（atm）	T_{cell}（K）	T_{fuel}^{in}（K）
		T_{air}^{in}（K）

8.2.4　电压

燃料电池在阳极和阴极之间会产生电动势，这是由式（7.1）所示的电化学反应所致，而且这也是燃料电池的电压输出来源。在正常操作范围内，启动损耗和浓度亏损要比欧姆损耗小得多，可以忽略不计，而 SOFC 堆栈的电压输出可以表示为

$$V_{out}=N\left[E^0+\frac{RT_{cell}}{2F}\ln\left(\frac{p_{H_2}p_{O_2}^{\frac{1}{2}}}{p_{H_2O}}\right)-R_{in}I\right] \tag{8.1}$$

式中，N 是电池堆栈中的电池数量，可以认为在没有损耗的情况下 $N=1$；在标准

状态下的燃料电池电动势是 $E^0=1.273-2.7645\times10^{-4}T_{cell}$；$p_{H_2}$、$p_{O_2}$和 p_{H_2O}是反应物的分压；而 R_{in}是燃料电池的内阻。

输出电压 V_{out}的动态反应会受到电流 I，温度 T_{cell}和 p_{H_2}、p_{O_2}、p_{H_2O}反应物分压的影响。

输出电压 V_{out}也会在扩散过程中受到动态影响和双电层电容的影响，正如第 7 章所讨论的那样。欧姆损耗和浓度亏损分别可以反映双电层电容和扩散过程对输出电压的影响情况。扩散过程和双电层电容对输出电压的影响时间非常短，大约在 $10^{-2}\sim10^{-1}$s。因此，这两种情况对电压的动态影响可以忽略不计，只有在稳态下需要特别考虑。

8.2.5　分压

反应物的分压直接影响输出电压。反过来，分压又受温度、流体输送过程和重整、转移以及电化学反应的影响。模拟结果表明，由流体动态、重整/转移和电化学反应引起的分压，反应速度快（在 $10^{-2}\sim10^{0}$s 的数量级），如第 7 章所述。因此，可以忽略这些动态变化，而只考虑它们的稳态行为。

重整和水煤气转移反应对燃料气体的组成起到了重要决定性作用。它们从甲烷和蒸汽中产生 H_2 和 CO：

重整：$$CH_4+H_2O \rightleftharpoons 3H_2+CO \tag{8.2}$$

转移：$$CO+H_2O \rightleftharpoons H_2+CO_2$$

当燃料进入到阳极通道，在阳极的催化剂（镍）和电池中的热量的条件下，发生重整和转移反应。沿流动方向，甲烷被逐渐消耗，此过程中产生氢。Campanari 和 Iora（2004）发现了稳态组成的曲线。

重整反应速率取决于燃料温度（Ahmed 和 Foger，2000），并且通常比流动速率的动态特性快。这意味着，甲烷是在堆栈内完全反应。蒸汽被重整和转移反应所消耗。同时，该电化学反应产生蒸汽。重整及转移反应产生氢。然后通过电化学反应消耗。消耗速率是由需求电流 I 确定的，如式（7.11）所示。

可以在质量守恒定律的基础上计算各组分的分压。详细推导过程已在第 7 章中给出。在稳态下，残留在流动通道各成分的摩尔数可以通过减去初始供给的摩尔消耗数来计算。各物质的分压，由留在流动通道的摩尔数来确定（见表 8.2）。

在表 8.2 中，Δ 是由转换反应所消耗的摩尔数，$\dot{Q}_{fuel}$和 $\dot{Q}_{air}$分别为燃料和空气的体积流动速率。

最终组成是由转移反应（认为是迅速达到平衡）确定的。平衡常数取决于燃料温度，其定义如下：

$$K_{eq}=\frac{p_{CO_2}p_{H_2}}{p_{CO}p_{H_2O}}=\exp\left(\frac{4276}{T_{fuel}}-3.961\right) \tag{8.3}$$

由转换反应消耗的摩尔数 Δ 可以由式（8.3）来求得：

表 8.2　在 SOFC 堆栈中从进口到出口的物质变化

	初始	整合	变化	回应	残余
CH_4	$P_{fuel}^{in}\chi_{CH_4}^{in}$	$-P_{fuel}^{in}\chi_{CH_4}^{in}$	-0	-0	0
H_2O	$P_{fuel}^{in}\chi_{H_2O}^{in}$	$-P_{fuel}^{in}\chi_{CH_4}^{in}$	$-\Delta$	$+\frac{I}{2F}RT_{fuel}\frac{1}{\dot{Q}_{fuel}}$	p_{H_2O}
H_2	$P_{fuel}^{in}\chi_{H_2}^{in}$	$+3P_{fuel}^{in}\chi_{CH_4}^{in}$	$+\Delta$	$-\frac{I}{2F}RT_{fuel}\frac{1}{\dot{Q}_{fuel}}$	p_{H_2}
CO	$P_{fuel}^{in}\chi_{CO}^{in}$	$+P_{fuel}^{in}\chi_{CH_4}^{in}$	$-\Delta$	-0	p_{CO}
CO_2	$P_{fuel}^{in}\chi_{CO_2}^{in}$	-0	$+\Delta$	-0	p_{CO_2}
O_2	$P_{air}^{bot}\chi_{O_2}^{in}$	-0	-0	$-\frac{I}{4F}RT_{air}\frac{1}{\dot{Q}_{air}}$	p_{O_2}

$$\Delta = \frac{K_{eq}(a_1 + a_3) + a_2 + a_4}{2(K_{eq} - 1)} - \frac{\sqrt{[K_{eq}(a_1 + a_3) + a_2 + a_4]^2 - 4(K_{eq} - 1)(K_{eq}a_1a_3 - a_2a_4)}}{2(K_{eq} - 1)} \tag{8.4}$$

式中

$$a_1 = P_{fuel}^{in}\chi_{H_2O}^{in} - P_{fuel}^{in}\chi_{CH_4}^{in} + \frac{I}{2F}\frac{RT_{fuel}}{\dot{Q}_{fuel}}$$

$$a_2 = P_{fuel}^{in}\chi_{H_2}^{in} + 3P_{fuel}^{in}\chi_{CH_4}^{in} - \frac{I}{2F}\frac{RT_{fuel}}{\dot{Q}_{fuel}}$$

$$a_3 = P_{fuel}^{in}\chi_{CO}^{in} + P_{fuel}^{in}\chi_{CH_4}^{in}$$

$$a_4 = P_{fuel}^{in}\chi_{CO_2}^{in}$$

使用上述方程，分压 p_{H_2O}、p_{H_2}、p_{CO}和 p_{CO_2}可由输入变量 p_{fuel}^{in}、I 以及初始组成 $\chi_{CH_4}^{in}$、$\chi_{H_2O}^{in}$、$\chi_{H_2}^{in}$、χ_{CO}^{in}和$\chi_{CO_2}^{in}$来计算。

8.2.6　流速

如第 7 章所述，燃料和气体流速也决定着反应物的组成，并且燃料和空气流速主要受 3 个因素影响：上游压力、下游压力和流动阻力。当出口压力固定在 1atm 时，此时流速由进气压力和流动阻力决定。

8.2.6.1　燃料流动速率

通常，燃料的流速在 1 ~ 10m/s 的范围。所以电池堆栈的停留时间为 0.1 ~ 1s。与缓慢的固相温度响应相比，快速流动的动态响应是很快的。此外，由于燃料和空气是可压缩流体，任何影响流体的干扰都可以导致高速度的短波移动。因此，有待开发一种流动过程的稳态模型。

在如图 8.3 所示的燃料流量通道中，关于燃料流体的稳态力平衡由以下公式表示：

$$\tau_{w,fuel}2\pi r_3L=(P_{fuel}^{in}-P_{fuel}^{exit})(4-\pi)r_3^2 \tag{8.5}$$

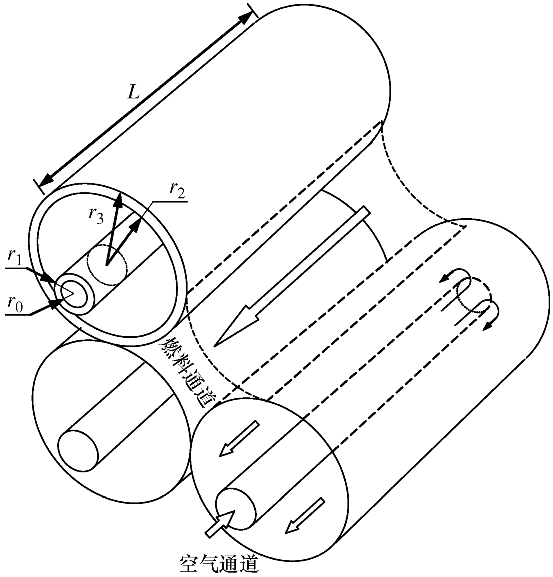


图 8.3　管式 SOFC 堆栈中燃料和空气的流动通道

式中，$\tau_{w,fuel}$是壁面切应力；r_3 是电池外管的半径；L 是流道的长度。

因此

$$\tau_{w,fuel}=\frac{(4-\pi)r_3}{2\pi}\frac{101325(P_{fuel}^{in}-P_{fuel}^{exit})}{L} \tag{8.6}$$

对于在特定通道的层流，摩擦系数 f 和雷诺数 Re 的数量是固定的。一个形状如图 8.3 所示的燃料通道，可以近似看成一个方管，因此其 $fRe=56.91$（John 和 Haberman，1988）。

由 f 和 Re 的定义可以得到：

$$fRe=\frac{\tau_{w,fuel}}{\frac{1}{8}\rho\,\bar{u}^2}\cdot\frac{\rho\,\bar{u}D_h}{\mu}=56.91 \tag{8.7}$$

式中，$D_h=\dfrac{4(4-\pi)r_3^2}{2\pi r_3}$表示燃料通道的水力直径。

处理式（8.7）中的燃料流速 $\bar{u}_{fuel}$并替换式（8.6）中的结果，可以得到体积流速的表达式是

$$\begin{aligned}\dot{Q}_{fuel}&=A_a\,\bar{u}_{fuel}=(4-\pi)r_3^2\cdot\frac{8}{56.91}\cdot\frac{\tau_{w,fuel}D_h}{\mu\,\bar{u}_{fuel}}\\&=101325\cdot\frac{8(4-\pi)^3r_3^4}{56.91\pi^2L}\cdot\frac{1}{\mu_{fuel}}\cdot(P_{fuel}^{in}-P_{fuel}^{exit})\end{aligned} \tag{8.8}$$

8.2.6.2　气体流速

与燃料流速不同，空气通过氧化铝注入管注入 SOFC 管底部。然后，燃料通过如图 8.3 所示的电池注入管和管之间的间隙进入阴极流道，并在其中进行流动。

气流速度可以达到 10m/s。因此，只需要考虑稳定状态。

从模型实验得到的仿真结果表明，如第 7 章所述，无论是在注入管还是阴极流道内，流动即层流（Qi 等人，2006）。因此，在注入管内，可以得到 $fRe=64$。以此类推，注入管内空气的流速为

$$\dot{Q}_{\text{air}}^{\text{inj}}=101325\cdot\frac{\pi r_0^4}{8L}\cdot\frac{1}{\mu_{\text{air}}}\cdot(P_{\text{air}}^{\text{in}}-P_{\text{air}}^{\text{bot}})\tag{8.9}$$

式中，$P_{\text{air}}^{\text{bot}}$是电池管底部的空气压力，这也是阴极流道的入口压力；r_0 为注入管的内半径。

阴极流道是一个环通道。空气流速由下式计算得到：

$$\dot{Q}_{\text{air}}=101325\cdot\frac{\pi r^*}{8L}\cdot\frac{1}{\mu_{\text{air}}}\cdot(P_{\text{air}}^{\text{bot}}-P_{\text{air}}^{\text{exit}})\tag{8.10}$$

式中

$$r^*=r_2^4-r_1^4-\frac{(r_2^2-r_1^2)^2}{\ln(r_2/r_1)}\tag{8.11}$$

式中，r_2 和 r_1 分别是电池管的内半径和注入管的外半径。

注入管内的空气继续流向阴极通道，因此流速保持不变：

$$\dot{Q}_{\text{air}}=\dot{Q}_{\text{air}}^{\text{inj}}\tag{8.12}$$

把式（8.9）和式（8.10）代入式（8.12），分析计算可得，电池管底部的空气压力是

$$P_{\text{air}}^{\text{bot}}=\frac{r_0^4\cdot P_{\text{air}}^{\text{in}}+r^*\cdot P_{\text{air}}^{\text{exit}}}{r_0^4+r^*}\tag{8.13}$$

8.2.7　温度

正如第 7 章所介绍的，温度在 SOFC 动态中起着重要作用。有 5 个不同的温度影响着电池，它们是气相燃料温度 T_{fuel}、注入管的空气温度 $T_{\text{air}}^{\text{inj}}$、阴极的环境温度 T_{air}、固相电池的温度 T_{cell}、注入管的温度 T_{tube}。

8.2.7.1　气相温度

燃料和空气都以 1~10m/s 的速度流动。它们的温度不会随着负荷、速度波动和固相温度的变化而有显著的反应。燃料和空气的比热相对比较小，响应速度非常快，响应时间在 10^{-1}s 左右。所以假定管内气体温度和入口温度相同：

$$\begin{gathered}T_{\text{fuel}}=T_{\text{fuel}}^{\text{in}}\\T_{\text{air}}=T_{\text{air}}^{\text{inj}}=T_{\text{air}}^{\text{in}}\end{gathered}\tag{8.14}$$

以第 7 章提出的详细模型为基础做的仿真验证了这个近似值。

8.2.7.2 固相温度

由于大的比热，温度对固体电池管和注入管的响应大约为 10^2s，正如第 7 章介绍的那样。详细的分析显示固相温度（特别是电池的温度 T_{cell}）控制着 SOFC 的传递响应。因此电池和注入管的温度需要做动态的仿真。

把电池管作为一个控制量。下面几项总结了能量在控制量中的输入和输出：

1）反应物释放化学能，JH，J 是进入电池的质量流，H 是生成焓；

2）电能用于供给外部负载 $V_{out}I$；

3）热的交换是通过电池管和周围的对流，$hA\ (T_w - T_f)$，这里 h 是传热系数；

4）热交换网络通过电池和注入管辐射交换，$\frac{\sigma}{R_{rad}}A(T_{cell}^4 - T_{tube}^4)$，这里 R_{rad} 是辐射热阻；

5）热流被重整反应吸收，$r_{ref}\Delta H$，这里 r_{ref} 是重整反应速率，ΔH 是反应热。

根据能量守恒定律，内能的改变在控制量范围内由下式给出：

$$\begin{aligned} m_{cell}C_{p,cell}\frac{dT_{cell}}{dt} = &\frac{I}{2F}\cdot H_{H_2}\big|_{T_{fuel}} + \frac{I}{4F}\cdot H_{O_2}\big|_{T_{air}} - \frac{I}{2F}\cdot H_{H_2O}\big|_{T_{cell}} - \\ &V_{out}\cdot I - h_a\cdot 2\pi r_3 L\cdot(T_{cell} - T_{fuel}) - h_c\cdot 2\pi r_2 L\cdot(T_{cell} - T_{air}) - \\ &\frac{\sigma}{R_{rad}}\cdot 2\pi r_2 L\cdot(T_{cell}^4 - T_{inj}^4) - \\ &A_r K_r p_{CH_4}^{\alpha} p_{H_2O}^{\beta}\cdot \exp\left(-\frac{E_r}{RT_{fuel}}\right)\cdot \Delta H\big|_{T_{fuel}} \end{aligned} \tag{8.15}$$

注入管在热传导过程中是一个内部环境。然而，由于它的高比热，所以它影响着电池的温度。注入管吸收的热是通过电池管的辐射和用强对流与周围环境进行的热交换：

$$\begin{aligned} m_{inj}C_{p,inj}\frac{dT_{inj}}{dt} = &\frac{\sigma}{R_{rad}}\cdot 2\pi r^2 L\cdot(T_{cell}^4 - T_{inj}^4) - h_c\cdot 2\pi r_1 L\cdot(T_{inj} - T_{air}) - \\ &h_t\cdot 2\pi r_0 L\cdot(T_{inj} - T_{air}) \end{aligned} \tag{8.16}$$

8.3 非线性状态空间模型

定义输入矢量 $\boldsymbol{u}$：

$$\boldsymbol{u} = [u_1 \quad u_2]^T = [P_{fuel}^{in} \quad P_{air}^{in}]^T \tag{8.17}$$

干扰矢量 $\boldsymbol{d}$：

$$\boldsymbol{d} = [d_1 \quad d_2 \quad d_3]^T = [I \quad T_{fuel}^{in} \quad T_{air}^{in}]^T \tag{8.18}$$

输出矢量 $\boldsymbol{y}$：

$$\boldsymbol{y} = [y_1 \quad y_2]^{\mathrm{T}} = [V_{\mathrm{out}} \quad T_{\mathrm{cell}}]^{\mathrm{T}} \tag{8.19}$$

状态矢量 $\boldsymbol{x}$：

$$\boldsymbol{x} = [x_1 \quad x_2]^{\mathrm{T}} = [T_{\mathrm{cell}} \quad T_{\mathrm{inj}}]^{\mathrm{T}} \tag{8.20}$$

在表 8.2 的表达式中代入式（8.8）、式（8.10）和式（8.13），然后将表 8.2 的分压替换到式（8.1）的电压表达式中，SOFC 的非线性状态空间模型可以描述如下：

状态：

$$\begin{aligned}\dot{x}_1 = \frac{1}{m_{\mathrm{cell}}C_{p,\mathrm{cell}}}\Bigg\{&\frac{1}{2F}\left(H_{\mathrm{H_2}} + \frac{1}{2}H_{\mathrm{O_2}} - H_{\mathrm{H_2O}}\right)d_1 - \\ &\left[E^0 + \frac{R}{2F}x_1 \ln\left(\frac{p_{\mathrm{H_2}}p_{\mathrm{O_2}}^{0.5}}{p_{\mathrm{H_2O}}}\right) - R_{\mathrm{in}}d_1\right]d_1 - \\ &h_{\mathrm{a}} \cdot 2\pi r_3 L \cdot (x_1 - d_2) - h_{\mathrm{c}} \cdot 2\pi r_2 L \cdot (x_1 - d_3) - \\ &A_r K_{\mathrm{r}} x_1^{(\alpha+\beta)} \cdot \chi_{\mathrm{CH_4}}^{\mathrm{in}\ \alpha} \chi_{\mathrm{H_2O}}^{\mathrm{in}\ \beta} \cdot \exp\left(-\frac{E_{\mathrm{r}}}{R \cdot d_2}\right) \cdot \Delta H\Bigg\}\end{aligned}$$

$$\begin{aligned}\dot{x}_2 = \frac{1}{m_{\mathrm{inj}}C_{p,\mathrm{inj}}}\Bigg\{&\frac{\sigma}{R_{\mathrm{rad}}} \cdot 2\pi r_2 L \cdot (x_1^4 - x_2^4) - \\ &-2\pi L \cdot (h_{\mathrm{c}} r_1 + h_{\mathrm{t}} r_0) \cdot (x_2 - d_3)\Bigg\}\end{aligned}$$

输出：

$$y_1 = E^0 + \frac{R}{2F}x_1 \ln\left(\frac{p_{\mathrm{H_2}}p_{\mathrm{O_2}}^{0.5}}{p_{\mathrm{H_2O}}}\right) - R_{\mathrm{in}}d_1$$

$$y_2 = x_1$$

式中

$$H_{\mathrm{H_2}} = -0.9959 \times 10^4 + 30.73 d_2$$

$$H_{\mathrm{O_2}} = -1.2290 \times 10^4 + 35.12 d_3$$

$$H_{\mathrm{H_2O}} = -25.790 \times 10^4 + 42.47 d_2$$

$$E^0 = 1.273 - 2.7645^{-4} x_1$$

$$p_{\mathrm{H_2}} = u_1(\chi_{\mathrm{H_2}}^{\mathrm{in}} + 3\chi_{\mathrm{CH_4}}^{\mathrm{in}}) + \Delta - 1.0955 \times 10^{-3}\frac{R}{2F} \cdot \frac{\mu_{\mathrm{fuel}}L}{r_3^4} \cdot \frac{d_1 d_2}{u_1 - P_{\mathrm{fuel}}^{\mathrm{exit}}}$$

$$p_{\mathrm{H_2O}} = u_1(\chi_{\mathrm{H_2O}}^{\mathrm{in}} - \chi_{\mathrm{CH_4}}^{\mathrm{in}}) - \Delta + 1.0955 \times 10^{-3}\frac{R}{2F} \cdot \frac{\mu_{\mathrm{fuel}}L}{r_3^4} \cdot \frac{d_1 d_2}{u_1 - P_{\mathrm{fuel}}^{\mathrm{exit}}}$$

$$\begin{aligned}p_{\mathrm{O_2}} = &\frac{0.21 r_0^4}{r_0^4 + r^*}u_2 + \frac{0.21 r^* P_{\mathrm{air}}^{\mathrm{exit}}}{r_0^4 + r^*} \\ &-2.5132 \times 10^{-5}\frac{R}{4F} \cdot \frac{\mu_{\mathrm{air}}L}{r_0^4} \cdot \frac{r_0^4 + r^*}{r^*} \cdot \frac{d_1 d_3}{u_1 - P_{\mathrm{air}}^{\mathrm{exit}}}\end{aligned}$$

$$\Delta H = -20.73 \times 10^4 - 19.41 d_2$$

$$R_{\mathrm{in}} = 2.94 \times 10^{-5}\exp\left(\frac{10350}{x_1}\right) \cdot \frac{0.12 \times 10^{-3}}{2\pi r_3}$$

或者用一个简写形式：

$$\boldsymbol{x} = f(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{d})$$
$$\boldsymbol{y} = g(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{d}) \tag{8.21}$$

8.4　仿真

8.4.1　验证

动态模型探讨了能够影响 SOFC 动态性能的物质机理，以便设计出控制方案。因为 SOFC 技术正在发展过程中，实验数据，特别是动态实验数据非常少，并且这些数据的可重复性差。因此，并不打算从实验数据入手直接修改模型。作为一个简化的模型，本章中的低阶状态空间模型是通过和其他基于文献中可以获得的详细模型仿真结果模拟验证的。

所给出方程中 SOFC 堆栈参数见表 8.3。由于电池是以串联的形式简单排列形成堆栈，因此只需在方程中研究一个电池单元即可。

表 8.3　SOFC 堆栈设计参数

符　　号	描述
几何参数	
$L = 1.5\text{m}$	电池长度
$r_3 = 11 \times 10^{-3}\text{m}$	外部电池管半径
$r_2 = 8.66 \times 10^{-3}\text{m}$	内部电池管半径
$r_1 = 6 \times 10^{-3}\text{m}$	外部注入管半径
$r_0 = 5 \times 10^{-3}\text{m}$	内部注入管半径
固体材料性质	
$\rho_{\text{cell}} = 4592\text{kg/m}^3$	SOFC 整体密度
$C_{p,\text{cell}} = 740\text{J/(kgK)}$	SOFC 整体比热
$\rho_{\text{e}} = 2.94 \times 10^{-5}\exp\left(\frac{10350}{T_{\text{cell}}}\right)\Omega\text{m}$	电解质电阻率
$\rho_{\text{tube}} = 3900\text{kg/m}^3$	注入管密度
$C_{p,\text{tube}} = 976.8 + 0.2409T\ \text{J/(kgK)}$	注入管比热
重整于转化	
$K_{\text{r}} = 8542$	重整反应速率系数
$E_{\text{r}} = 95\text{kJ/mol}$	重整反应活化能
$\alpha = 0.85$	重整阶数
$\beta = -0.35$	重整阶数
$Ar = 1 \times 10^{-4}\text{m}^2$	重整反应面积
热传递	
$h_{\text{a}} = 8.744\text{Wm}^2/\text{K}$	阳极处对流传热系数
$h_{\text{c}} = 63.83\text{Wm}^2/\text{K}$	阴极处对流传热系数
$h_{\text{t}} = 79.57\text{Wm}^2/\text{K}$	注入管对流传热系数

（续）

符　　号	描述
$R_{rad}=4.69$	辐射传热阻力
注入成分	
$\chi_{CH_4}^{in}=0.173$	注入 CH_4 摩尔分数
$\chi_{H_2O}^{in}=0.284$	注入 H_2O 摩尔分数
$\chi_{H_2}^{in}=0.258$	注入 H_2 摩尔分数
$\chi_{CO}^{in}=0.057$	注入 CO 摩尔分数
$\chi_{CO_2}^{in}=0.228$	注入 CO_2 摩尔分数

在图 8.4 中模拟电压动态响应从 300A 的电流阶跃变化到 500A，与文献中其他可用的仿真结果进行了比较。在它们的模拟中，Hall 和 Colclaser（1999）只考虑了堆栈的温度动态。Padullés 等人（2000）在假设温度保持不变的情况下模拟了阻流压力的影响。在 Padullés 等人工作的基础上，Sedghisigarchi 和 Feliachi（2004a）总结了温度动态的影响。Iora 等人（2005）建立了一个更详细的模型，通过考虑流体动态，重整反应和热转换等因素。现有结果已经表明基于本章中建立的阶跃响应模型和文献中相关模型一致。

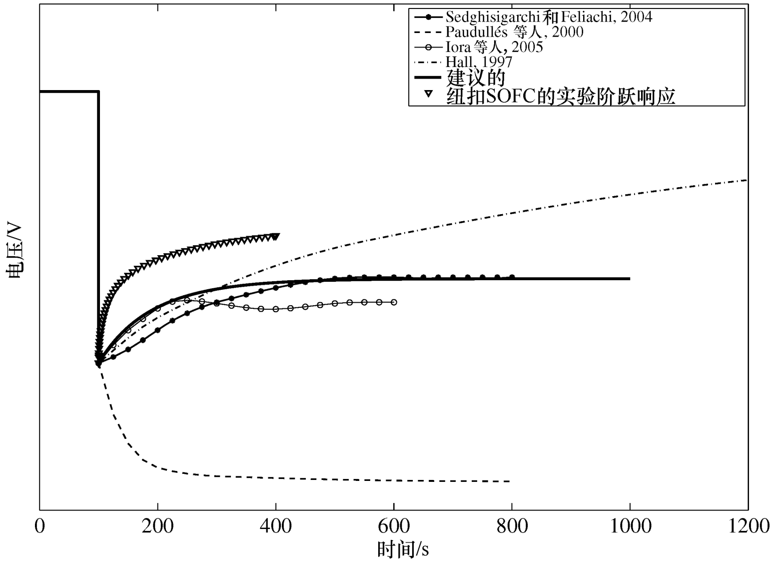


图 8.4　当前需求阶跃改变响应的电压堆栈比较

阶跃模型和实验室中纽扣电池装置测试结果性质相似。尽管实验数据不能完全比较，因为纽扣电池的构造和管结构不同。实验结果反映了 SOFC 模型阶跃响应的整体发展趋势：当电流增大，由于本身电阻的存在，电压迅速降低；电流因为电阻消耗而加热电池，使电池温度逐渐升高。如式（7.4）所示，EMF 取决于电池温度，因此输出电压逐渐升高直到终值，如图 8.4 所示。模型的趋势和实验的结果一致。

由于文献中缺少实验数据和对燃料/空气入口流动波动动态反应的模型结果，模拟电压响应的趋势与第 7 章中建立的详细模型的模拟结果相比较，已经证明由于入口波动引起的动态响应和第 7 章模拟的详细模型的动态响应一致。

8.4.2 输入阶跃响应

如图 8.5 所示，当燃料的上游压力 P_{fuel}^{in} 从 1.0008atm 增加到 1.0016atm，合成气燃料的流动速率迅速增加一倍。这意味着更多的 CH_4 在堆栈内转化成了 H_2。相应的，氢气的分压力 p_{H_2} 增加，同时输出电压 V_{out} 迅速增加。更多的热量被从电池释放以维持重整反应的进行，因此，电池温度 T_{cell} 轻微但逐渐下降，因为它本身具有很大的热容量。这个结果使 V_{out} 略微增加，与 T_{cell} 的反应同步。

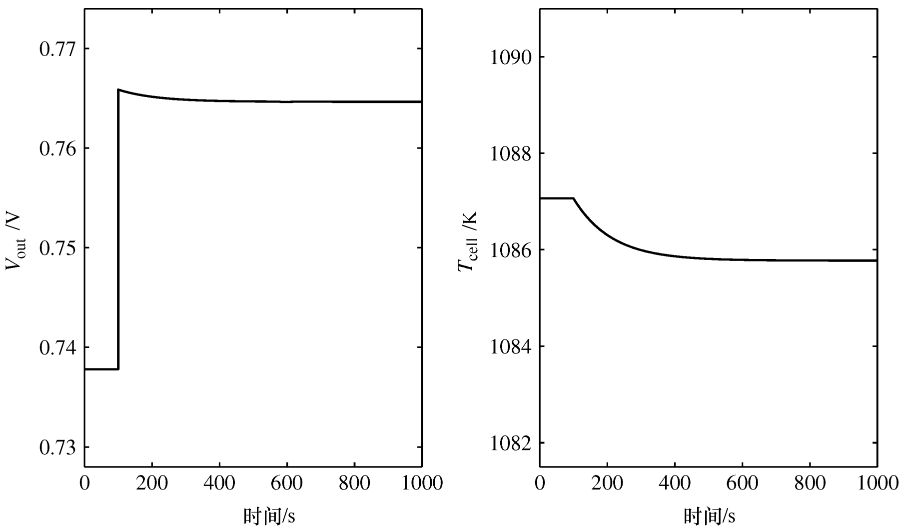


图 8.5 由于 u_1，上游燃料气压从 1.0008atm 增加到 1.0016atm 时的输出阶跃响应

和燃料输入压力增加的影响类似，当上游空气压力 P_{air}^{in} 从 1.01atm 增加到 1.02atm 时，电池堆栈内氧气分压 p_{O_2}、管内的空气流动速率和输出电压同时增加。阶跃响应如图 8.6 所示。

传真图直接揭示了入口燃料压力 p_{fuel}^{in} 阶跃变化，同样对输出电压 V_{out} 有重要影响。它对电池温度 T_{cell} 的影响是微小、缓慢的，和慢一阶反应相似，时间恒定在 150s 左右。空气入口压力 P_{air}^{in} 的影响和燃料入口压力的影响类似。

8.4.3 干扰阶跃响应

堆栈的阶跃对干扰的影响来源于电源需求、入口燃料和空气温度，已经通过仿真研究过。

当外部电流 I 需求从 300A 增加到 500A，堆积的反应如图 8.7 所示。可以明显看出，输出电压 V_{out} 对电流 I 阶跃变化的响应是迅速和明显的。所需电流 I 对电池温度 T_{cell} 也有巨大的影响。由于电池管巨大的热容量，电池温度的阶跃响应和一阶

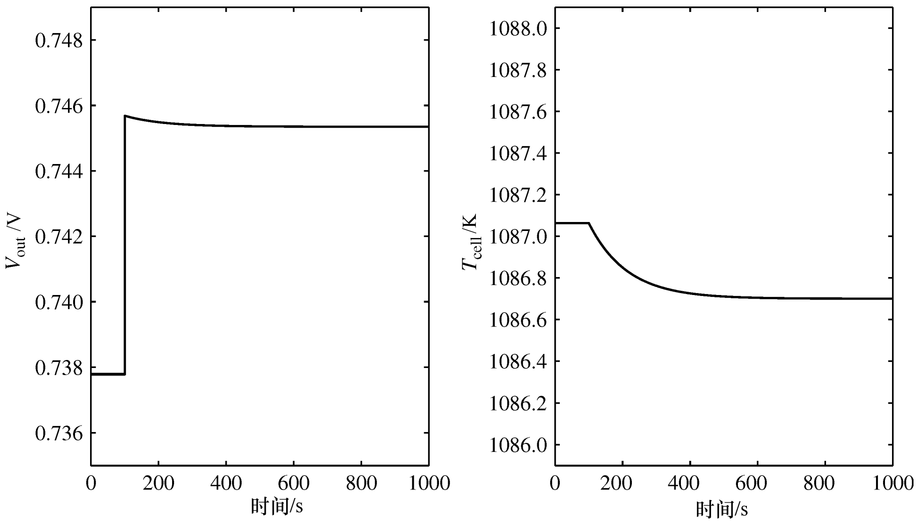


图 8.6　由于 u_2，上游空气气压从 1.01atm 增加到 1.02atm 时的输出阶跃响应

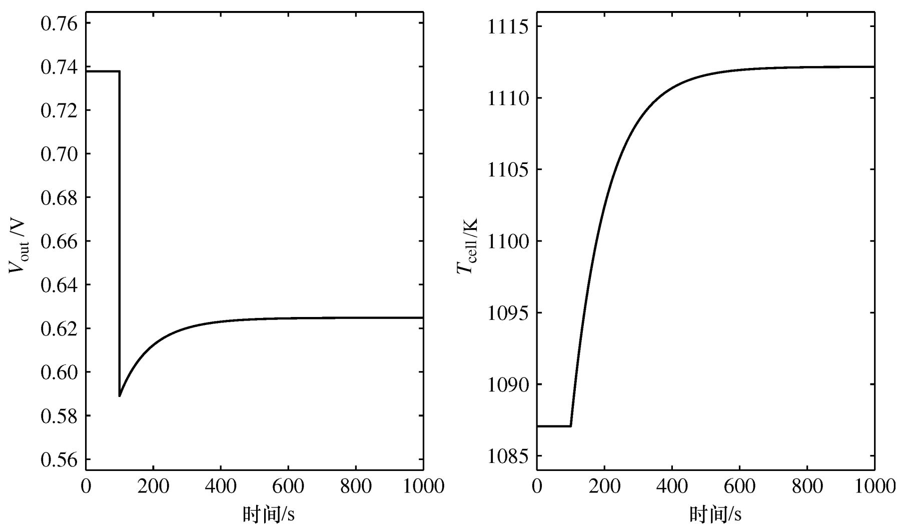


图 8.7　由于 d_1，当前电流 I 从 300A 增加到 500A 时的输出阶跃响应

响应保持时间在 150s 左右。

由于电流 I 的增加，电压由于电池的内阻而迅速降低，同时更多的氢气和氧气被消耗。所以氢气分压 p_{H_2} 和氧气分压 p_{O_2} 降低，水的分压 p_{H_2O} 增大。结果就是输出电压 V_{out} 在短时间内下降。内阻消耗的能量随电流 I 的增加而增大。这部分能量使电池升温，相应的 T_{cell} 增加。所以输出电压 V_{out} 在最初的迅速下降之后随着电池温度 T_{cell} 的升高慢慢恢复。

当燃料入口温度 T_{fuel}^{in} 从 823K 增加到 873K 时，转移反应平衡被破坏，氢气分压

p_{H_2}降低，所以输出电压 V_{out}迅速降低。同时，T_{fuel}^{in}的升高降低了燃料和电池的温差。然后 T_{cell}增加保持时间在 150s 左右。输出电压 V_{out}和电池温度 T_{cell}对燃料入口温度 T_{fuel}^{in}的阶跃变化的阶跃响应如图 8.8 所示。

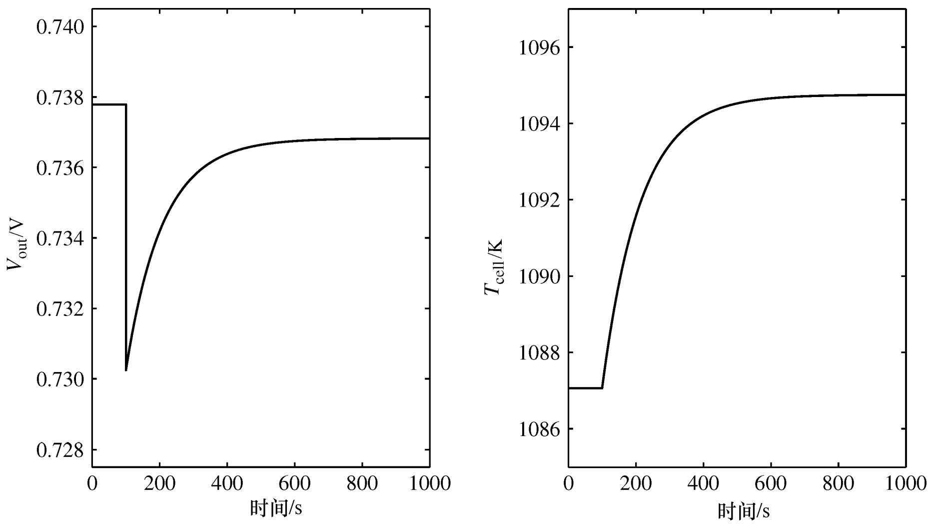


图 8.8　由于 d_2，燃料入口温度 T_{fuel}^{in}从 823K 增加到 873K 时的输出阶跃响应

空气入口温度 T_{air}^{in}干扰的影响更直接、更显著，当 T_{air}^{in}从 1104K 增加到 1154K 时，这使得电池加热，所以 T_{cell}逐渐增加。与此同时，随着 T_{cell}逐渐增加，输出电压 V_{out}也增加，如图 8.9 所示。

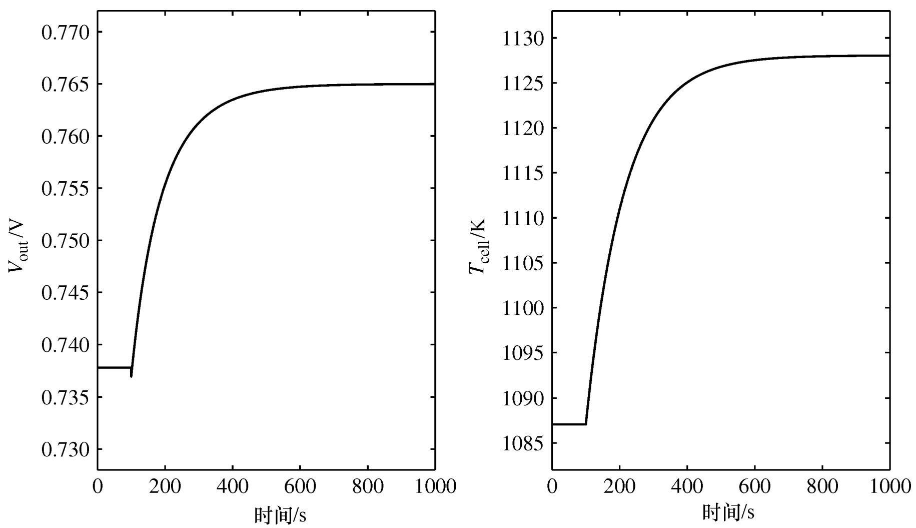


图 8.9　由于 d_3，空气入口温度 T_{air}^{in}从 1104K 增加到 1154K 时的输出阶跃响应

仿真图显示 V_{out} 和 T_{cell} 对于 T_{air}^{in} 的阶跃变化响应明显，该响应类似于具有大约 150s 时间常数的一阶响应。

所有前面所述动态仿真对输入和干扰变化的反应和第 7 章中建立的详细 28 阶 SOFC 仿真结果一致。这表明二阶模型可以近似获得 SOFC 堆栈的主要动态特征。

8.5　注解和参考文献

降低复杂的集总参数动态模型对控制和监控设计是有用的，也是广大读者对控制领域所关心的。读者可以在下面找到关于集总参数模型的资料：Achenbach（1994，1995）、Padullés 等人（2000b）、Zhu 和 Tomsovic（2001）、Sedghisigarchi 和 Feliachi（2004a）、Xue 等人（2005）、Gemmen 和 Johnson（2005）以及 Iora 等人（2005）的文献。

第 9 章　管式 SOFC 的动态建模与控制：系统辨识方法

9.1　简介

当第一性原理建模为 SOFC 过程中的短暂的行为提供了详细的细节时，这些模型的主要的缺点是它们的复杂性。这种复杂性使得实时控制的应用显得很困难。另一方面，利用数据驱动系统辨识是一种与控制相关的方法，并且能够大大降低模型的复杂度，但是也只能得到一种经验性的过程的输入和输出的关系。这种从系统辨识中得到的模型应当具有比第一性原理更少的参数，而且具有更简单的模型结构。这些模型的缺点是较差的推断能力。然而，为了得到在控制和监视方面的实时应用，经验模型一直是最常用的。

本章的主要目的是证明在第 3 章和第 4 章中系统辨识方法在第 7 章中用一种详细的第一性原理模型代表的 SOFC 中的应用。第 7 章中得到了几种简单的经验模型，在本章中将会比较各种经验模型和得到非线性的系统辨识方法。在这些被辨识的模型的基础上，商用的 PID 控制器在最初的第一性原理模型的环境中进行设计和测试。

9.2　系统辨识

9.2.1　变量选择

第一性原理模型的状态空间表示具有 16 个输入、28 个状态和 17 个输出。在这项研究中，电池的温度仅作为一个状态，并被指定作为一个输出。各个输入和输出的变量见表 9.1。在这项研究中，感兴趣的输出变量是电池的输出电压和电池温度。并不是所有的输入都能够用于系统的控制。例如，负载的电阻、混合物和燃料/空气的入口温度是不能或者不易被操纵的。这些变量可以被认为是扰动变量。剩下的输入变量是入口的燃料气体和喷入的空气的压力和速度。压力和速度是容易通过阀门和泵来控制的。因此，这 4 个变量被选为可操纵变量的可能的候选变量。

表 9.1　输入和输出变量

输 入 变 量	输 出 变 量
R_{load}	V_{out}，I_{out}，T_{cell}
P_{fuel}^{in}，T_{fuel}^{in}，u_{fuel}^{in}	P_{fuel}，T_{fuel}，u_{air}

（续）

输入变量	输出变量
$\chi_{CH_4}^{in}$，$\chi_{H_2O}^{in}$，$\chi_{H_2}^{in}$，χ_{CO}^{in}，$\chi_{CO_2}^{in}$	χ_{CH_4}，χ_{H_2O}，χ_{H_2}，χ_{CO}，χ_{CO_2}
P_{air}^{in}，T_{air}^{in}，u_{air}^{in}，$\chi_{O_2}^{in}$	P_{air}，T_{air}，u_{air}，χ_{O_2}
P_{inj}^{in}，T_{inj}^{in}，u_{inj}^{in}	P_{inj}，T_{inj}，u_{inj}

9.2.2 阶跃响应测试

作为系统辨识的第一步，执行阶跃响应测试通常用来得到近似的过程的参数。增益的大小用来提供对于输出的稳态影响的输入变化信息。如果影响变为可以忽略的话，这个变量可能不能作为有效的可操纵的变量。虽然这种说法不完全正确，但是这种方法却经常被用来快速筛选潜在的调节变量。第二个在阶跃测试中可以被快速确定的变量是过程的时间常数。这个信息提供了过程动态的深入信息，并且可以帮助设计输入信号和决定可操纵变量的有效性。在一个最简单的阶跃测试过程中，每一个阶跃变化都被当成一个输入变量，其他的所有输入都保持常数。所有输出变量的响应都被记录下来。在一个典型的过阻尼线性系统中，输出响应如图 9.1 所示。输入变量的阶跃变化在时间 1 时从 0 变化到 1，输出的响应如图 9.1 中的实线所示。这个过程中的增益可以从输出变化和输入变化的比值来确定，其中，输出变化指的是新的和最初的稳态值的区别值。一阶系统的死区时间可以从阶跃响应上输入信号变化开始的时间和输出信号变化开始的时间的差值直接得到。

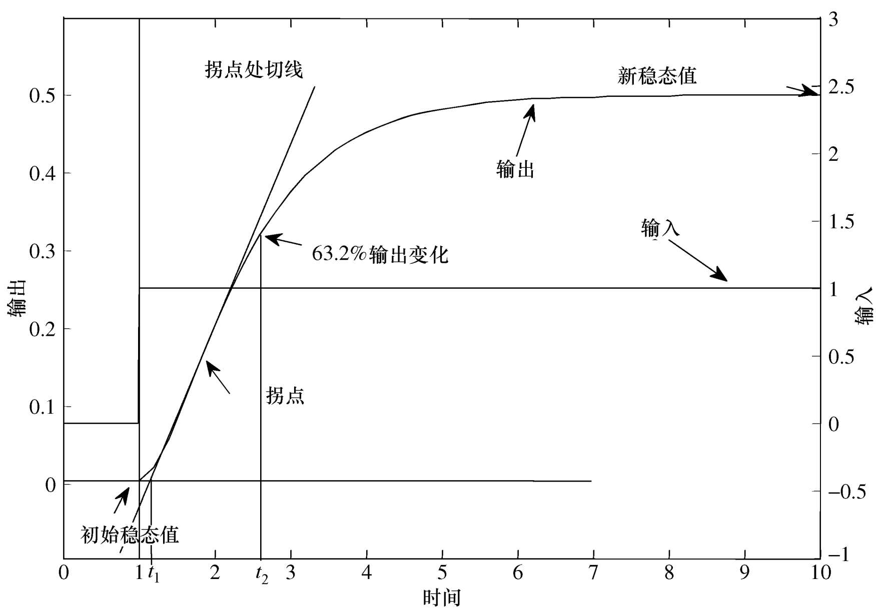


图 9.1 从阶跃响应中计算增益、时间延迟和时间常数

一个高阶系统的死区时间是由一个近似得到的。死区时间的起点是当输入变化开始的时候。然后，输出响应的拐点首次被发现，然后通过这个点可以画出切线。在这个切线上可以发现一个点，这个点的 y 轴的值是初始的稳态值，响应的 x 轴的值是时间延迟的终点。在本例中，死区时间是（t_1-1）。时间常数是整个输出改变值的 63.2%。在本例中，时间常数是（t_2-t_1）。

由于 SOFC 是一个非线性的过程，所以线性模型辨识只能应用在围绕选定的工作点的局部动态。在本测试中，考虑 3 个分别围绕工作点 $R_{load}=2\Omega$、4Ω、8Ω 的负载的线性动态模型的辨识。这三个工作条件分别被归类为低负载（LL）、额定负载（NL）和高负载（HL）。

图 9.2 说明了在工作点 $R_{load}=2\Omega$ 的电池温度的阶跃响应，此时输入变量，燃料入口气体的压力从 1atm 变到 1.3atm，采样时间为 1000s。从图 9.2 中可以看出，初始的稳态值是 1080.8K。新的稳态值大约为 1076.2K。因此，增益 k 是

$$k=\frac{1076.2\text{K}-1080.8\text{K}}{1.3\text{atm}-1\text{atm}}=-15.3\text{Katm}^{-1}$$

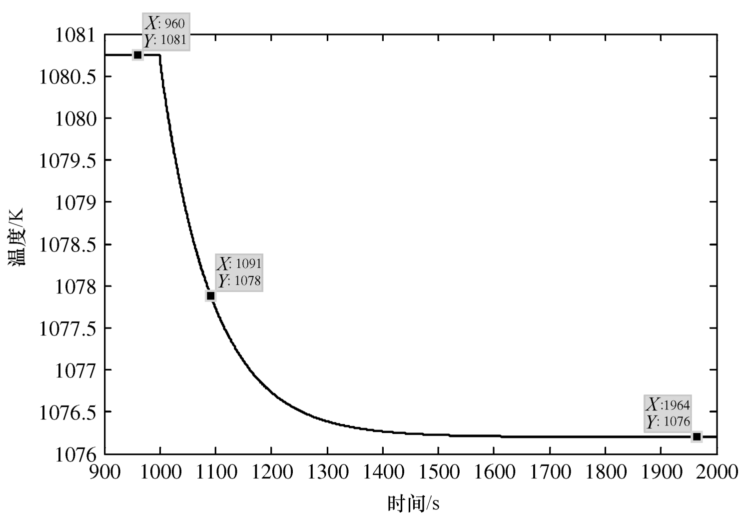


图 9.2　当燃料气压从 1～1.3atm/1000s 时电池温度的阶跃响应

这个过程的死区时间可以忽略，因为输出的变化正好在 1000s。确定输出的 63.2% 的变化可以找到这个过程的时间常数。具体方法是：首先得到 0.632 ×（1076.2K − 1080.8K）= −2.91K。然后，在这个点的输出值是 1080.8K +（−2.91K）= 1077.9K。最后，通过找到输出值为 1077.9K 的点，得到时间常数为 91s。

采用同样的方法，每一个感兴趣的输入和输出对在 3 个不同的工作点的增益、死区时间和时间常数都可以确定。表 9.2 显示了时间常数。在一些有"—"标记的位置表示相应的过程不能近似为一种一阶加时间延迟模型。这种类型的过程将在后面讨论。

表 9.2　对测试流程中 3 个不同操作方式下的时间常数　（单位：s）

	T_{cell}			电压		
	LL（2Ω）	HL（8Ω）	NL（4Ω）	LL（2Ω）	HL（8Ω）	NL（4Ω）
U_{fuel}	93	97	110	—	—	—
U_{inj}	104	104	110	93	105	100
P_{fuel}	91	93	92	—	—	—
P_{inj}	103	105	114	98	113	110

表 9.3 和表 9.4 给出了所有对的增益和死区时间（时间延迟），从中可以看出，所有对的时间延迟比时间常数要小得多。而增益的大小是为系统辨识方法选择目标变量的一个重要的因素。对于使用电池温度作为输出变量的那些 I/O 对来说，入口燃料气体的速度和压力是更显著的输入变量。当电压为输出变量时，只有入口燃料气体的压力对输出有很大的影响。在本章的剩余部分，两个 I/O 对（一个是入口燃料气体的速度和电池温度；另一个是入口燃料气体的压力和电池电压）被用来说明系统辨识的运用。这两对被标记为模型 1 和模型 2。

表 9.3　3 个不同操作模式下测试流程增益

	T_{cell}/K			电压/V		
	LL（2Ω）	HL（8Ω）	NL（4Ω）	LL（2Ω）	HL（8Ω）	NL（4Ω）
U_{fuel}	−13.38	12.8	−413.02	-9.72×10^{-3}	1.5×10^{-3}	-5.48×10^{-3}
U_{inj}	0.0390	0.0983	0.0784	2.74×10^{-5}	1.12×10^{-5}	3.42×10^{-5}
P_{fuel}	−15.89	14.7	−15.15	0.0554	0.027	0.0418
P_{inj}	0.463	1.17	0.874	3.33×10^{-4}	1.67×10^{-4}	3.6×10^{-4}

表 9.4　在 3 个不同操作模式下测试过程的时间延迟　（单位：s）

	T_{cell}			电压		
	LL（2Ω）	HL（8Ω）	NL（4Ω）	LL（2Ω）	HL（8Ω）	NL（4Ω）
U_{fuel}	0	0	0	0	0	0
U_{inj}	8	10	12	9	10	13
P_{fuel}	0	0	0	0	0	0
P_{inj}	9	9	8	10	8	8

9.2.3　非典型性阶跃响应

入口燃料气体的速度和压力阶跃变化时的电压响应不能近似为一阶加时间延迟模型。图 9.3 显示了当入口燃料气体的压力从 1atm 变化到 1.3atm 时的电压的阶跃响应。可以看出，当输入变化时，电压在很短的时间内有一个非常快速的动态变化，使得所有的动态被压缩成一条直线，接着是一个典型的一阶阶跃响应。

这个快速的动态放大了之后如图 9.4 所示。从图 9.3 可以看出，垂直线首先扩大到一个逆响应后随后有一个过冲现象，电压首先降到 0.0147V。如果通过输入阶

跃变化的量进行划分，这种“反增益”是 -0.049V atm^{-1}。注意到，从阶跃响应测试的总的增益是 0.0418V atm^{-1}，这个较大的逆响应是非常明显的。之后的过冲是 6.4%，是没有逆响应显著的。最初的动态的整个持续时间是 0.02s。为了捕捉这个动态，采样的时间应当小于 0.002s。然而，接下来的慢的动态不需要这么快的采样速度。因此，使用一般的采样速度的辨识方法的话，这个过程并不能被完全辨识。不过，用一个慢速动态过程的采样速度去应用于系统辨识的时候，这种慢速的动态过程可以被捕捉，但是快速的动态过程则不可以。

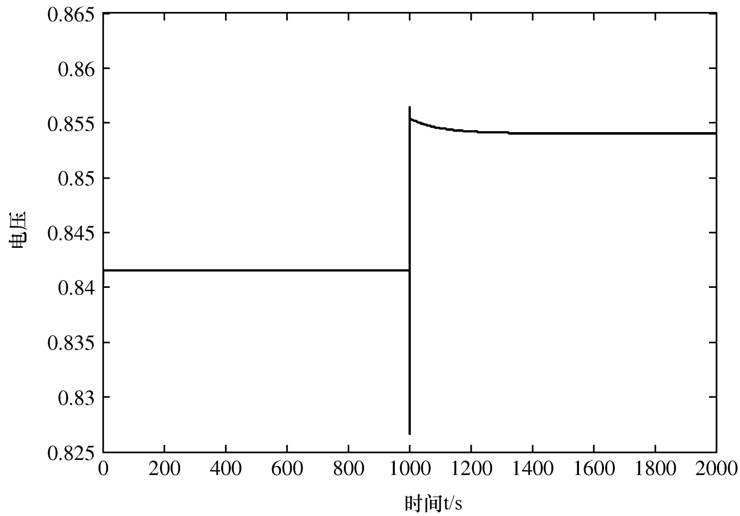


图 9.3　当燃料气压从 1 变化到 1.3atm/1000s 时电压的阶跃响应

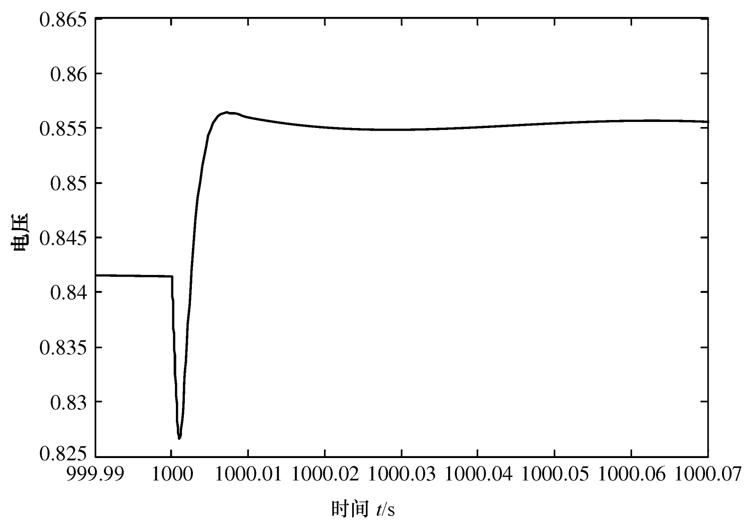


图 9.4　图 9.3 中的放大阶跃响应

9.2.4　输入设计

线性系统辨识将开展两个选择的 I/O 对。对于这两个模型的输入和输出信息显示在表 9.5 中。

表 9.5　考虑系统辨识的过程

输　入		输　出
模型 1	U_{fuel}	T_{cell}
模型 2	P_{fuel}	电压

该系统应被充分地由输入以执行系统身份激发。在本研究中使用的输入序列是随机二进制信号（RBS），其确定局部线性模型。在 MATLAB® 中的 RBS 生成命令是

$$\text{input} = \text{idinput}\ (N, '\text{rbs}',\ \text{band},\ \text{level})$$

式中，input 是输入变量的名称；N 表示将生成的输入采样的系数；band 表示输入信号频带的；level 包括在输入信号的两个开关电平之间。

频带通常由归一化频率指定。下限通常为零，而上限是最大的期望的频率的奈奎斯特（Nyquist）频率的一部分。上限的频带可以计算为

$$\text{upper bound} = \frac{kT_s}{\tau\pi}$$

式中，k 是一个常数，通常为 2 和 3（可以更大）；T_s 是数据采集系统的采样时间；τ 是该过程的时间常数。

在本研究中使用的 k 值是 2。采样时间被选择为 1/10 的每个进程主时间常数。根据表 9.2，所有进程的主时间常数大约为 100s，因此，在本研究的采样时间设定为 10s。以模型 2 为例，输入信号的频率的上限频带是

$$\text{upper bound} = \frac{2\times 10\text{s}}{100\text{s}\times\pi} = 0.064$$

根据表 9.2 可知，感兴趣的所有过程具有相似的时间常数，频带固定所有的输入信号，这是 0 ~ 0.064。二进制输入激励为 U_{fuel} 的幅度是 0.89ms^{-1} 和 0.96ms^{-1}，P_{fuel} 为 0.95atm 和 1.05atm。所产生的总的 10s 采样间隔的输入的 1000 点和相应的 RBS 输入序列生成命令是

$$U_\text{fuel_rbs} = \text{idinput}(1000, '\text{rbs}', [0\ 0.064], [0.89\ 0.96])$$

$$P_\text{fuel_rbs} = \text{idinput}(1000, '\text{rbs}', [0\ 0.064], [0.95\ 1.05])$$

使用第一性原理的动态模型进行仿真。4 个干扰源被添加到仿真中。第 1 个干扰来自燃料入口温度，这个是以 5℃标准偏差的白噪声通过一阶滤波器 $\frac{1}{10s+1}$ 产生的干扰。第 2 个干扰来自注入的空气温度，以 2℃标准偏差的白噪声通过一阶滤波器 $\frac{1}{10s+1}$ 产生的干扰。第 3 个干扰来自注入空气压力，这个是通过以 0.1atm 标准

偏差的白噪声到一阶滤波器$\frac{1}{s+1}$产生的。第4个干扰来自电池温度输出测量，是以0.1℃标准偏差的白噪声通过一阶滤波器$\frac{1}{s+1}$产生的。零阶保持器被用于前面的仿真操纵变量。

9.2.5　线性系统辨识

9.2.5.1　数据预处理

设计输入序列之后，仿真运行10000s。如前面所描述的输入和输出数据被每10s记录一次。系统辨识程序，对输入和输出数据进行处理。系统辨识模型1的例子将在后面详细说明。

输入和输出数据绘制在图9.5中。总仿真时间为10000s，采样时间间隔为10s，采集样本的总数为1000。使用MATLAB®系统辨识工具箱，一个数据对象构成如下：

$$zT = \mathrm{iddata}\ (T, U_\mathrm{fuel}, 1)$$

式中，T是电池温度响应；数据矢量U_fuel是燃料速度输入，这是通过零阶保持U_fuel_rbs后获得的；zT值是数据对象；1表示该时间单元以采样为单位。

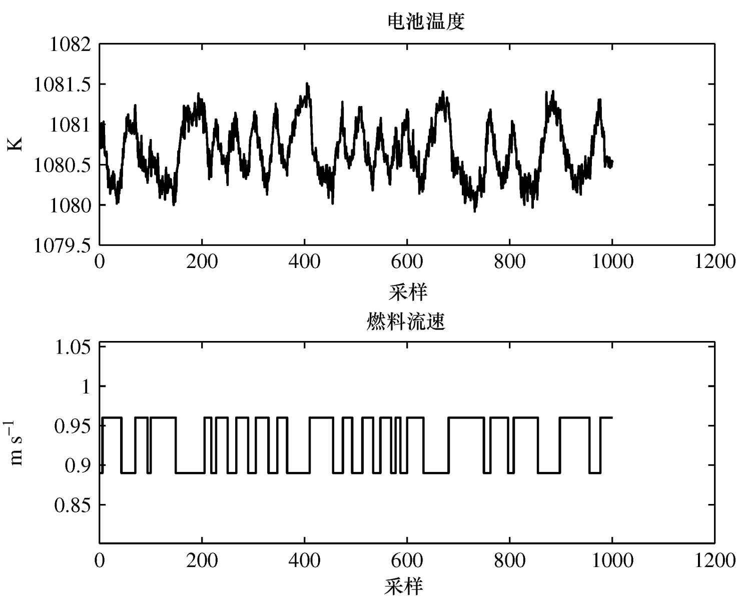


图9.5　输入和输出数据（模型1）

总数据集分为两部分：辨识数据集和验证数据集。辨识数据集用来估计模型中的参数，验证数据集用于交叉验证所辨识的模型。在这项研究中，通过下面的命令700个输入和输出数据用于辨识，剩余的300个样本用于交叉验证：

$$zTe = zT\ (1:700)$$

$$zTv = zT\ (701: \text{end})$$

zTe 中的数据将被用于辨识，zTv 中的数据用于交叉验证。

输入和输出变量都在执行任何系统辨识之前使用 MATLAB®命令 dtrend：

$$zTe = \text{dtrend}\ (zTe)$$

$$zTv = \text{dtrend}\ (zTv)$$

初审估算可以是互相关函数分析，可以从快速粗糙的脉冲响应估算和时间延迟得到。脉冲响应系数可以估算如下：

$$\text{cra}\ (zTe)$$

如图 9.6 所示，点画线包络代表在零处置信区间为 99%，即在包络区内任意处可视为零。估算给出过程中时间延迟的信息。从图中可看出，在置信区间以外的第一个脉冲响应点是采样时间 1。这意味着系统有一个样本延迟，主要是由零阶保持引起的。

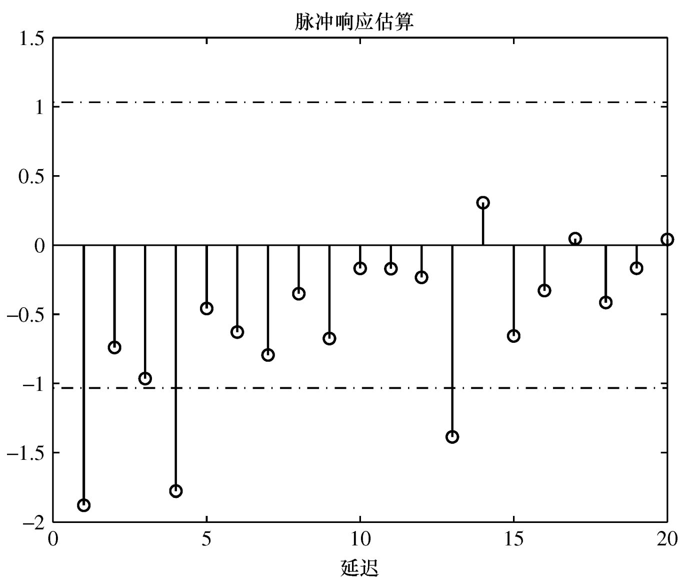


图 9.6　脉冲响应估算（模型 1）

9.2.5.2　ARX 模型

因为它们很简单，人们尝试了 ARX 模型的辨识。延时期限固定为 1。分母的不同次序和分子方面被认为是从 1 开始并且取决于交叉验证的结果逐渐增大。

例如，要确定一阶 ARX 模型，下面的 MATLAB®功能会被执行：

$$\text{ARX111} = \text{arx}\ (zTe,\ [1\ 1\ 1])$$

为了检验 ARX111，一个基于使用交叉验证数据集 zTv 的辨识模型的模拟响应，与实际响应相比较，下面的 MATLAB®功能：

$$\text{compare}\ (\text{ARX111},\ zTv)$$

的比较如图 9.7 所示，这表明一个差的预测性能的 ARX111。ARX111 的质量还通

过残差检验反映出来：

resid（ARX111，*zTv*）

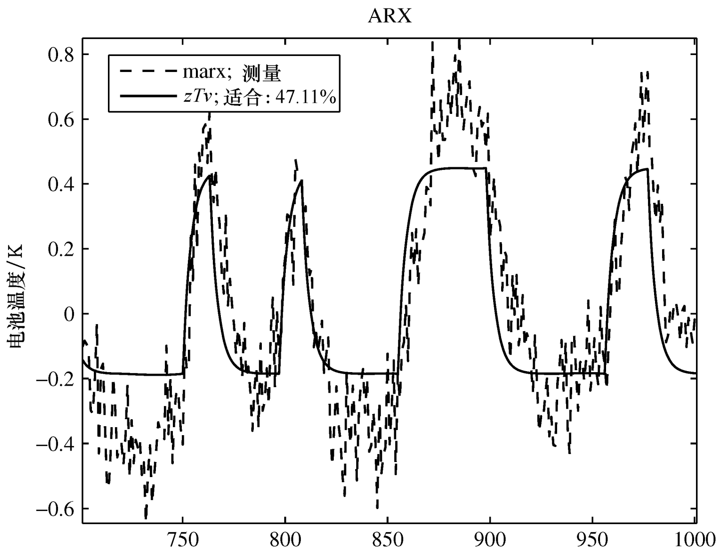


图 9.7　ARX111 预测与来自验证数据集的实际温度响应（模型 1）

残差的试验结果示于图 9.8，其中上图显示的是自相关检测结果，而下图显示的是互相关测试结果。两面的包络区显示在零处的置信区间为 99%。ARX111 模型互相关测试明显是失败的。超过 1% 以上的点位于置信区间之外。自相关测试也出现了问题或有可能出现问题。

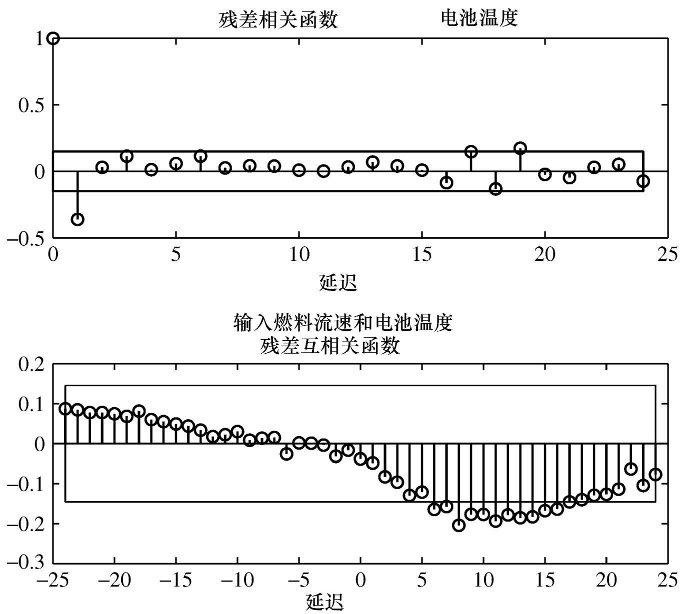


图 9.8　基于验证数据集的 ARX111 残差测试（模型 1）

因此，交叉验证结果表明，该一阶 ARX 模型在残留测试中是失败的。通过逐渐增加的次序，其性能得到了改善。例如，一个三阶 ARX 模型可以通过如下估算：

ARX331 = arx（zTe，[3 3 1]）

或者，可先尝试估算：

ARX321 = arx（zTe，[3 2 1]）

如果没有显著改善的话，就将分子增加一阶，然后增加分母的阶数。作为分母的确定模型阶，增加分母阶数是为了使系统阶数更高。

为了验证数据集，从第 3 阶 ARX 得到的预测响应模型与图 9.9 的实际响应相比较。图 9.10 显示了残差测试明确地传递。这些结果验证了该结论，即高阶 ARX 模型可以适当地与其他结构的模型近似。

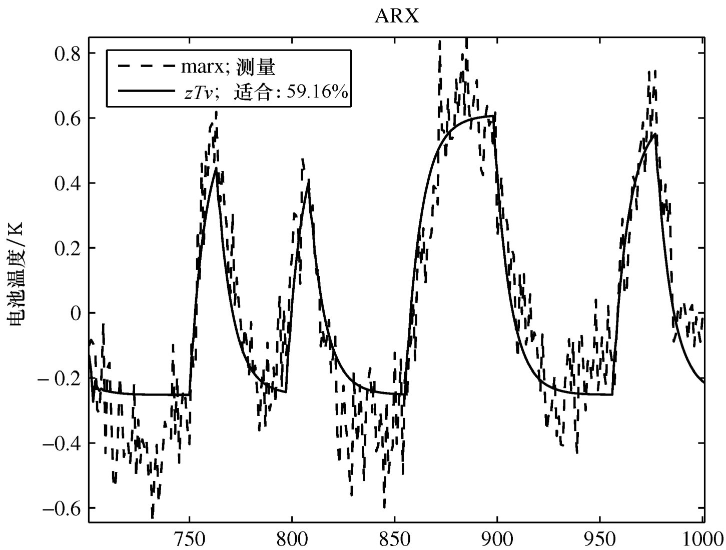


图 9.9　ARX331 预测与来自验证数据集的实际温度响应（模型 1）

对于三阶 ARX 模型的估算差分方程给定为

$$(1-0.2775(\pm.03662)z^{-1}-0.261(\pm.03585)z^{-2}-0.2363(\pm.03354)z^{-3})y_t$$
$$=(-1.63(\pm.313)z^{-1}-0.4484(\pm.4361)z^{-2}-0.6938(\pm.3418)z^{-3})u_t+e_t$$

其中括号括起来的是估算模型参数的标准偏差。在传递函数的形式下，所辨识的模式是

$$\frac{y_t}{u_t}=\frac{-1.63z^{-1}-0.4484z^{-2}-0.6938z^{-3}}{1-0.2775z^{-1}-0.261z^{-2}-0.2363z^{-3}}u_t$$
$$+\frac{1}{1-0.2775z^{-1}-0.261z^{-2}-0.2363z^{-3}}e_t$$

经验传递函数的博德图可以直接从数据以及通过以下功能先前确定的 ARX 模型估计：

mspa = spa（zTe）

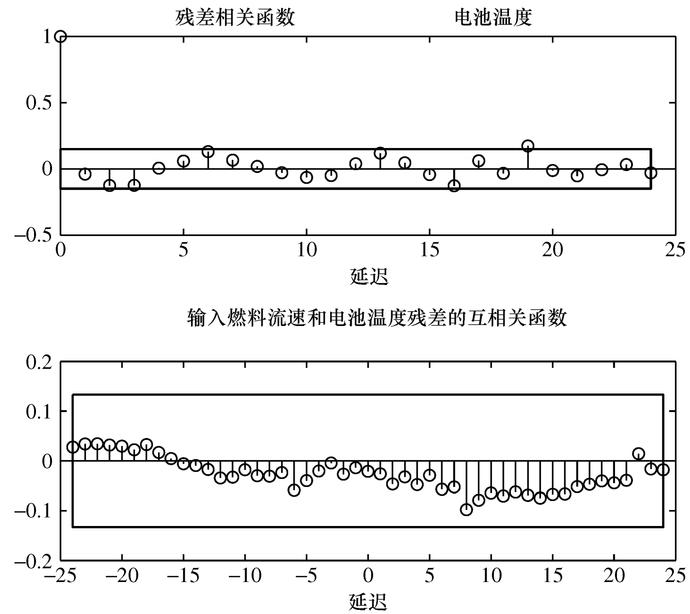


图 9.10　基于验证数据集的 ARX331 残差测试（模型 1）

bodeplot（marx331，mspa）

其结果示于图 9.11。可以看出，该 ARX331 模型从数据的总体趋势中直接估算出来的博德图捕获博德图。一般情况下，从数据直接估计博德图嘈杂较高，但往往偏差较少，而博德图从估计的参数化模型转换是平滑的，但更趋于偏颇。因此，这是一个比较好的方式来验证是否估计的参数化模型有偏差。

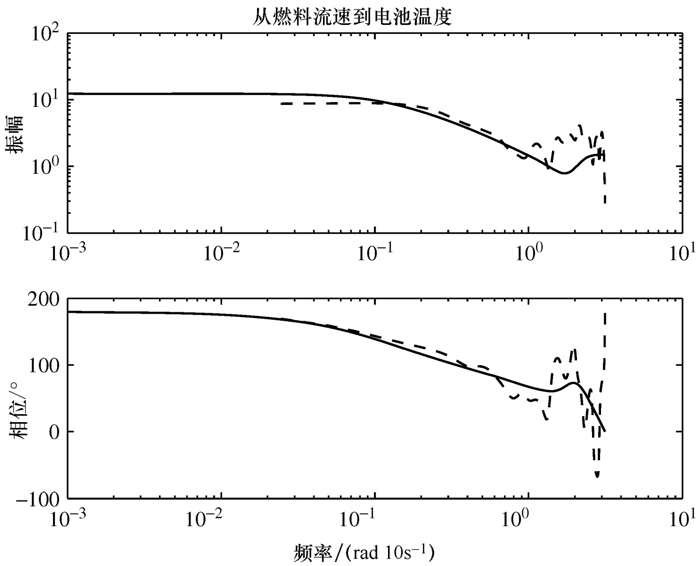


图 9.11　ARX331 模型博德图和基于频谱分析的经验传递函数估计（模型 1）

9.2.5.3　ARMAX 模型

该 ARMAX 模型在考虑之列。一阶 ARMAX 模型的辨识第一次尝试型号：

$$\text{ARMAX1111} = \text{armax}\ (zTe,\ [1\ 1\ 1\ 1])$$

这给出了确定模型的差分方程：

$$(1-0.9012z^{-1})y_t = -1.343z^{-1}u_t + (1-0.7008z^{-1})e_t$$

传递函数形式为

$$\frac{y_t}{u_t} = \frac{-1.343z^{-1}}{1-0.9012z^{-1}}u_t + \frac{1-0.7008z^{-1}}{1-0.9012z^{-1}}e_t$$

使用交叉验证数据集模拟的和实际的响应的比较如图 9.12 所示，残差测试都显示在图 9.13 中。与 ARX 模型残差测试类似 ARMAX 模型残差测试中还应该注意自相关和互相关测试。可以看到，在该测试中模型是成功的（在图 9.9 中示出 59.16%），一阶 ARMAX 模型略优于三阶 ARX 模型，虽然改善的效果在统计学中可能不是很显著。

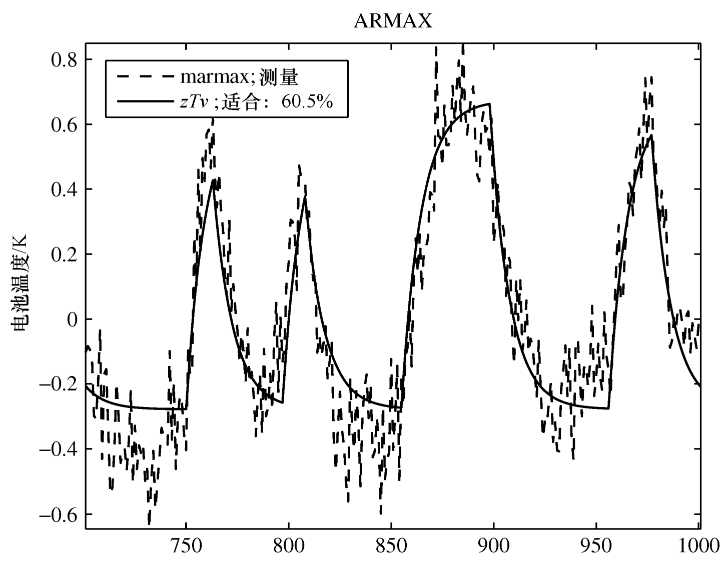


图 9.12　ARMAX1111 预测与来自验证数据集的实际温度响应（模型 1）

9.2.5.4　OE 模型

一阶 OE（输出误差）模型也试图通过如下辨识：

$$\text{OE111} = \text{OE}\ (zTe,\ [1\ 1\ 1])$$

估计模型的离散时间传递函数是

$$y_t = \frac{-1.296z^{-1}}{1-0.9071z^{-1}}u_t + e_t$$

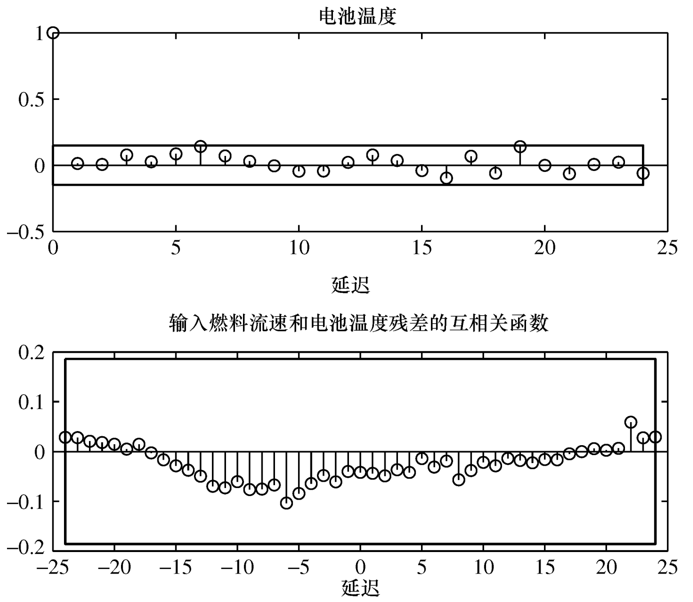


图 9.13　基于验证数据集的 ARMAX1111 残差测试（模型 1）

使用交叉验证数据集模拟和实际响应的比较示于图 9.14，残余的测试都显示在图 9.15。对于 OE 模型，自相关不必通过由于 OE 模型辨识算法的独立参数化性质的测试。唯一的互相关测试应予以考虑。可以看出，在残留测试中一阶 OE 模型是成功的。比较第一阶 OE 模型（见图 9.14 的 60.49%）与一阶 ARMAX 模型，它们有一个非常相似的性能，但参考模型给出了一个稀疏的模型结构。

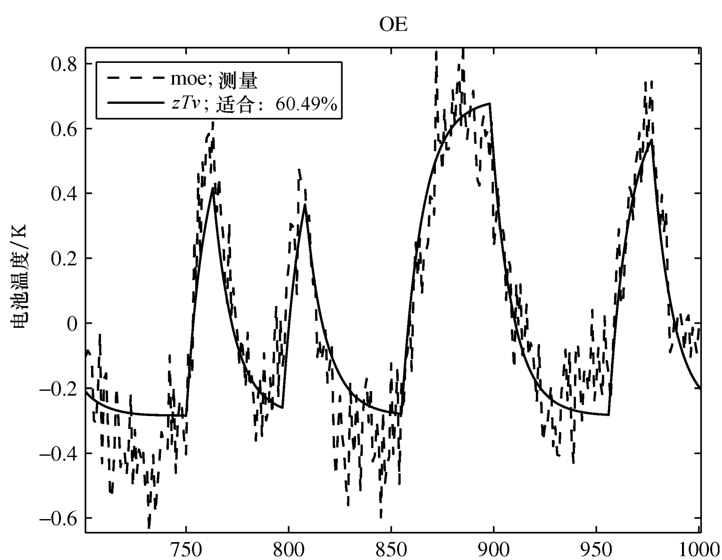


图 9.14　OE111 预测与来自验证数据集的实际温度响应（模型 1）

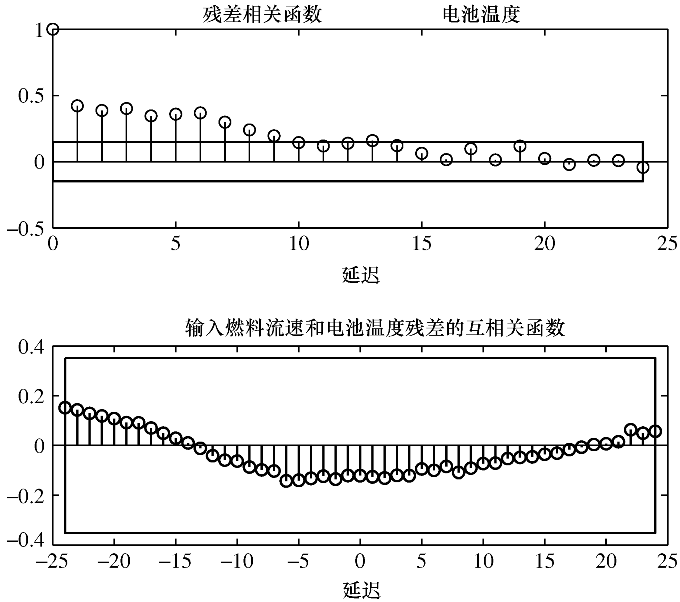


图 9.15　基于验证数据集的 OE111 残差测试（模型 1）

9.2.5.5　BJ 模型

一阶 BJ（Box - Jenkins）模型也已考虑使用以下函数：

估计 BJ 模型的传递函数是

$$y_t = \frac{-1.341z^{-1}}{1-0.9014z^{-1}}u_t + \frac{1-0.6947z^{-1}}{1-0.8969z^{-1}}e_t$$

使用交叉验证数据模拟的和实际的响应的比较如图 9.16 所示，残差测试都显示在图 9.17。对于一个 BJ 模型残差测试，应该看双方的自相关和互相关测试。人们可以看到，这个一阶 BJ 模型在任何残差测试中都是成功的。比较一阶 BJ 模型的拟合百分比（60.51%，见图 9.16）与一阶 ARMAX 模型和 OE 模型，它们有非常相似的性能，但 BJ 模型提供了完整的过程和干扰模型。

9.2.5.6　状态空间模型

通过子空间辨识一阶状态空间模型还通过以下功能考虑：

$$\text{SS1} = \text{n4sid}\ (zTe,\ 1)$$

估计状态空间模型是

$$x_{t+1} = 0.8962x_t - 0.1981u_t + 0.0325e_t$$

$$y_t = 6.9266z_t + e_t$$

使用交叉验证数据模拟的和实际的响应的比较如图 9.18 所示，残差测试都显示在图 9.19。对于状态空间模型的残差测试，应再次检查两个自相关和互相关测试。人们可以看到，这个一阶状态空间模型在任何残差测试中都是成功的。比较一

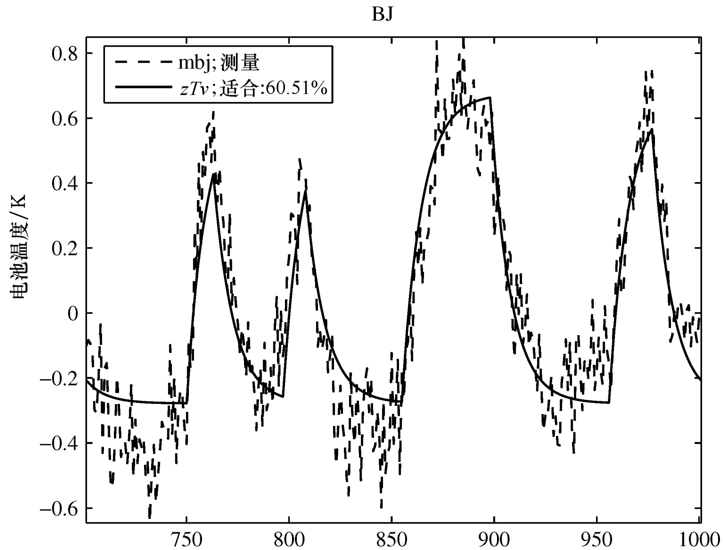


图 9.16　BJ11111 预测与来自验证数据集的实际温度响应（模型 1）

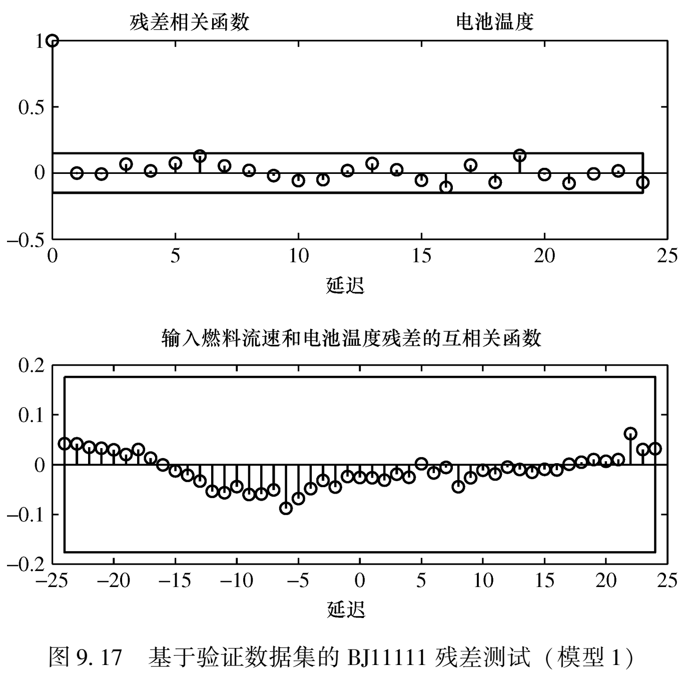


图 9.17　基于验证数据集的 BJ11111 残差测试（模型 1）

阶状态空间模型（60.3%，见图 9.18）与一阶 BJ 模型，ARMAX 模型和 OE 模型是比例合适的，它们又有着非常相似的性能，但是一种状态空间模型可以方便用于一些现代控制设计方法。

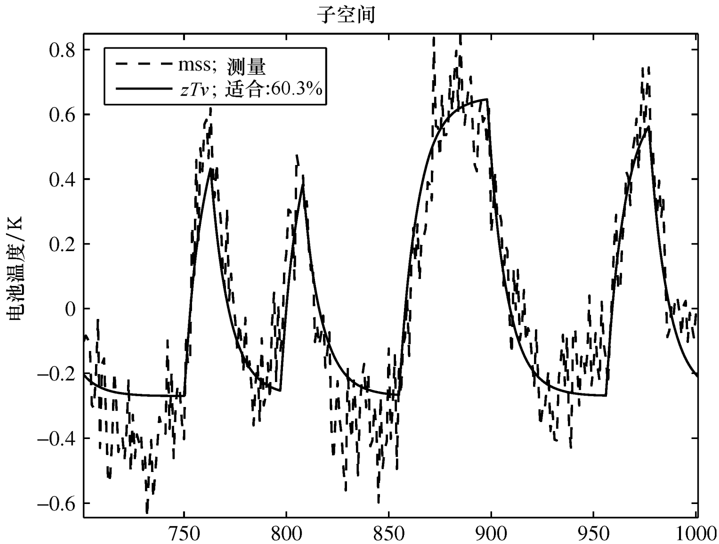


图 9.18　SS1 预测与来自验证数据集的实际温度响应（模型 1）

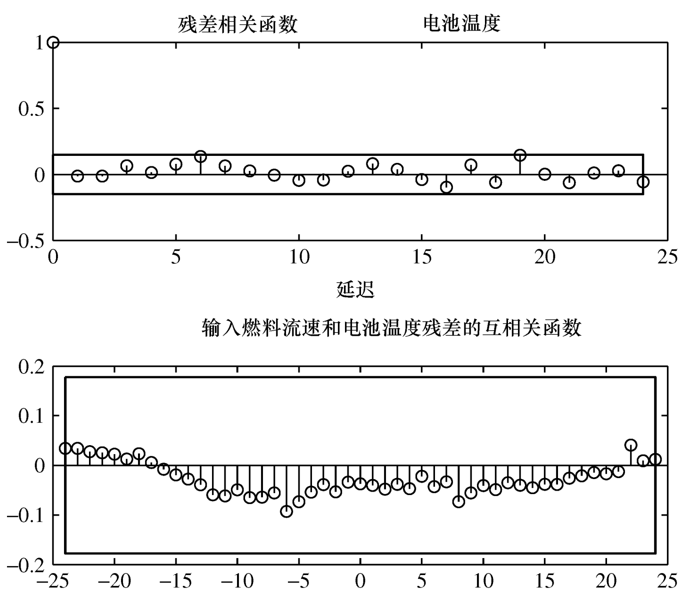


图 9.19　基于验证数据集的 SS1 残差测试（模型 1）

9.2.5.7　从输入燃料压力模型辨识电压输出

按照相同的程序，人们能够辨识线性模型（模型 2）与所述入口的燃料压力 P_{fuel}作为输入而电压作为输出。输入和输出数据如图 9.20 所示。首先，时间延迟可以通过相关分析估计从脉冲来确定反应，如图 9.21 所示。一个采样的时间延迟

可以从脉冲响应中清楚地看到。这种延迟是由于引入零阶保持而引起的。

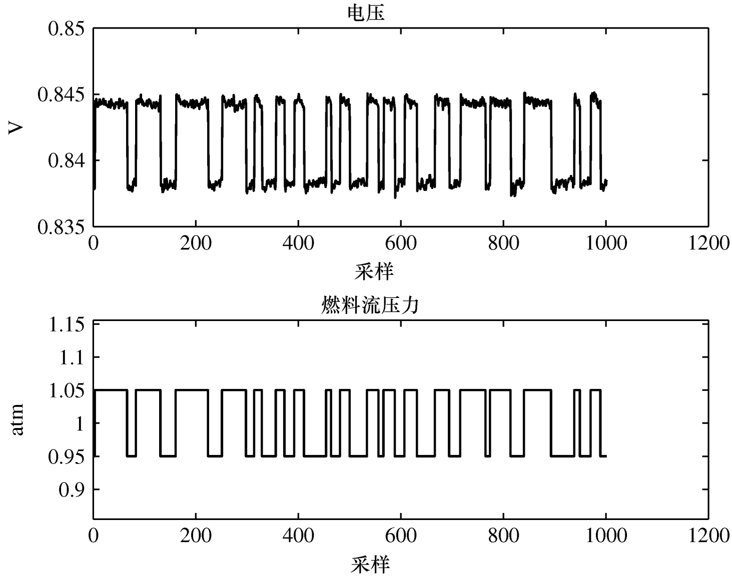


图 9. 20　输入和输出数据（模型 2）

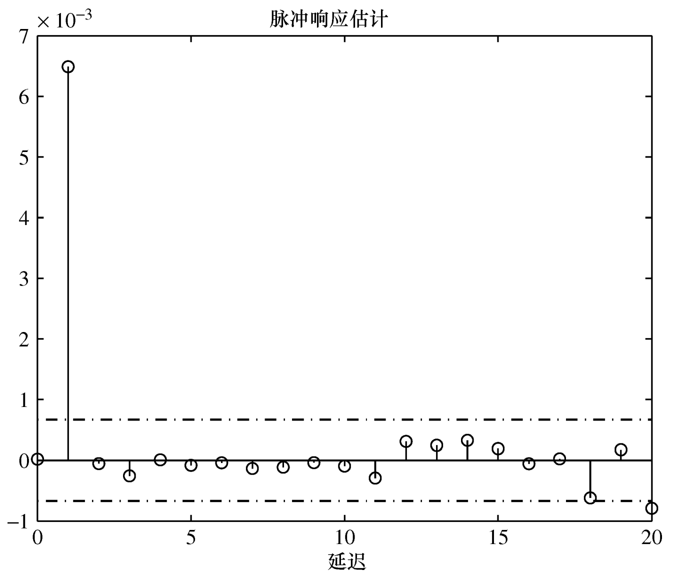


图 9. 21　脉冲响应估计（模型 2）

类似的情况已经在模型 1 中出现过了，人们一直尝试采用一阶 ARX 的辨识，ARMAX、OE、BJ 和 N4SID 模型也有尝试，但它们都没有通过残差测试。举一个例子，使用交叉验证的数据集的模拟和实际的比较使用所确定的参考模型的响应如

图 9. 22 所示。虽然合适的比例是高的（91. 39%），但它未能捕捉每次跳跃后的动态。残差测试结果为参考模型如图 9. 23 所示，在那里它可以清楚地看出，该模型中的互相关测试是失败的。

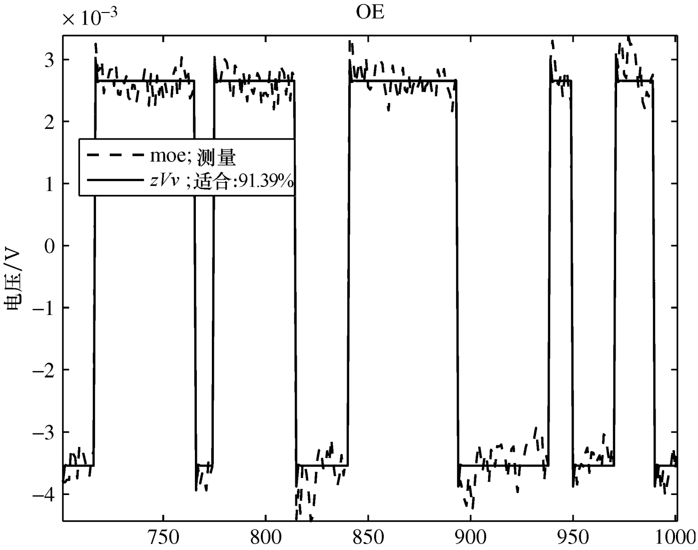


图 9. 22　OE111 预测与来自验证数据集的实际温度响应（模型 2）

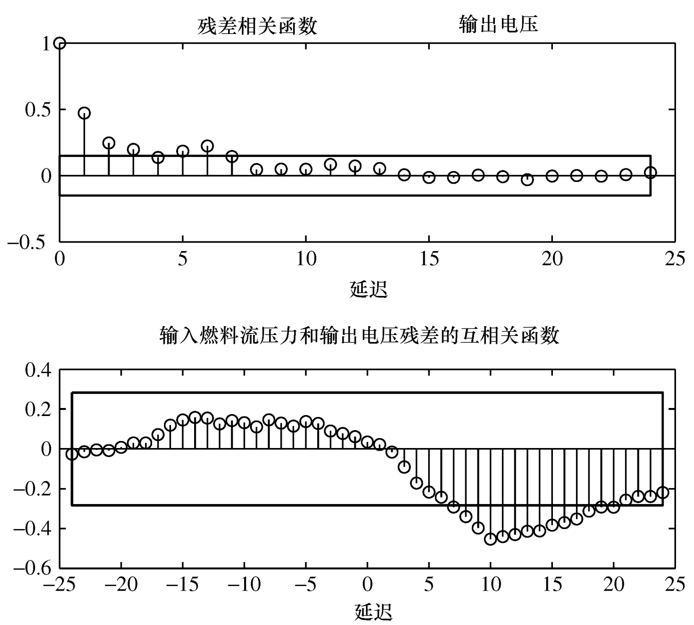


图 9. 23　基于验证数据集的 OE111 残差测试（模型 2）

类似的结果为 BJ 模型得到，如图 9. 24 和图 9. 25 所示，其中据观察，一阶 BJ 模型也会失败。

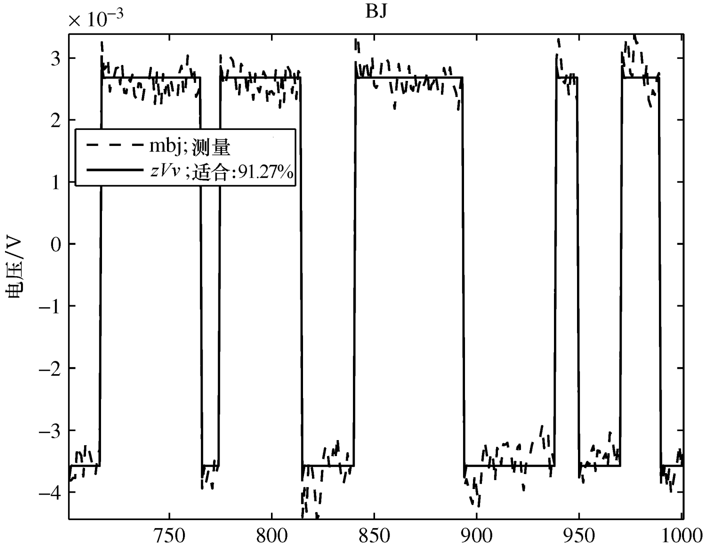


图 9. 24 BJ11111 预测与来自验证数据集的实际温度响应（模型 2）

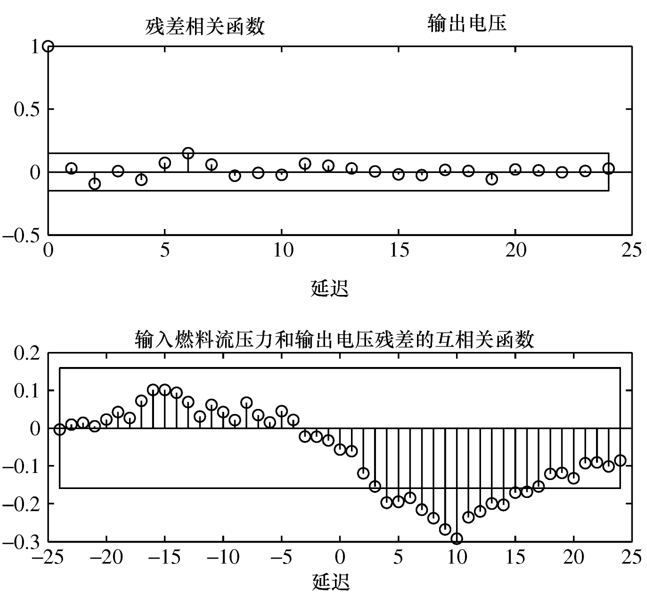


图 9. 25 基于验证数据集的 BJ11111 残差测试（模型 2）

接着，人们对二阶模型的辨识进行了尝试。不复述细节，估算 BJ 模型呈现在下面的传递函数中：

$$y_t = \frac{0.0648z^{-1} - 0.05901z^{-2}}{1 - 0.8976z^{-1} - 0.00373z^{-2}}u_t + \frac{1 - 0.9354z^{-1} + 0.05099z^{-2}}{1 - 1.25z^{-1} + 0.3602z^{-2}}e_t$$

验证结果为 BJ 模型，如图 9.26 和图 9.27 所示，其中一个可以看出，该模型能够捕捉到完整的动态。该模型在任何残差测试中都是成功的。所确定的二阶 BJ 博德图模型与经验传递函数估计的博德图相比，基于频谱分析的方法，并且将结果显示在图 9.28，这表明良好的一致性。所辨识的 BJ 模型的阶跃响应示于图 9.29。比较图 9.29 与图 9.30 所示的原始步骤的测试可以看到所确定的模型捕捉的整体变化，但没能把握住速率逆响应。这是预期的，如 1s 用于辨识采样速率太慢，不足以捕获 0.02s 内完成的逆反应。

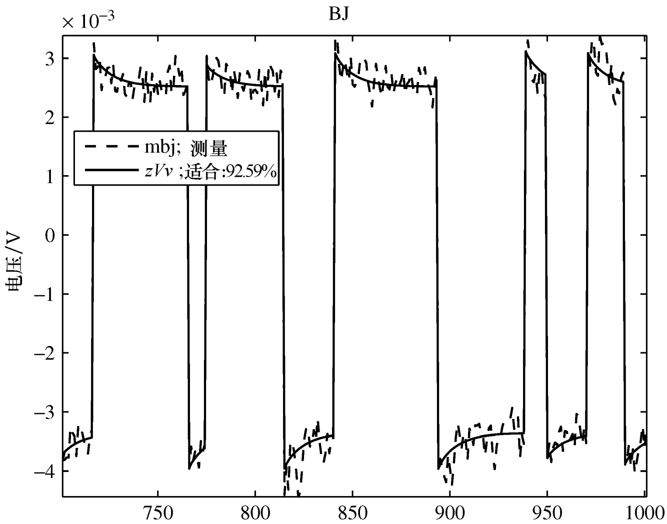


图 9.26　BJ22221 预测与来自验证数据集的实际温度响应（模型 2）

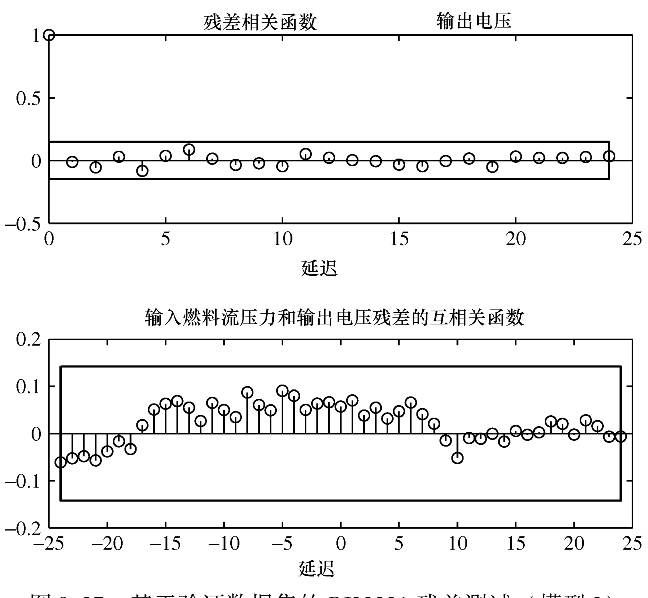


图 9.27　基于验证数据集的 BJ22221 残差测试（模型 2）

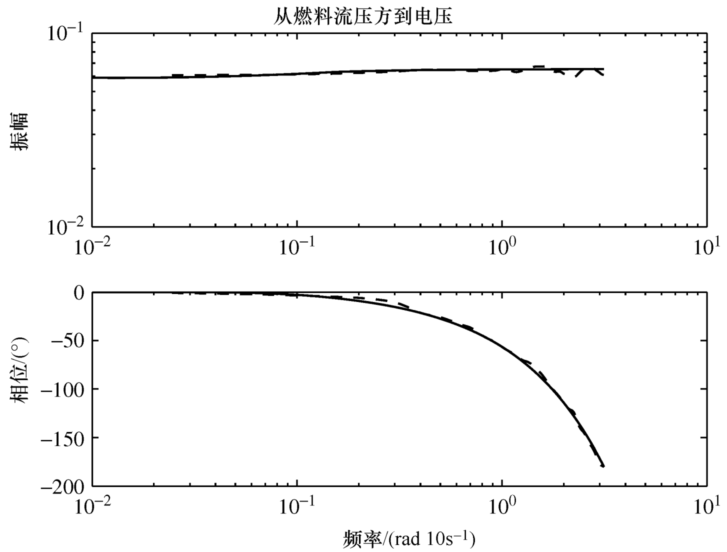


图 9.28 BJ22221 模型博德图和基于频谱分析的经验传递函数估计（模型 2）

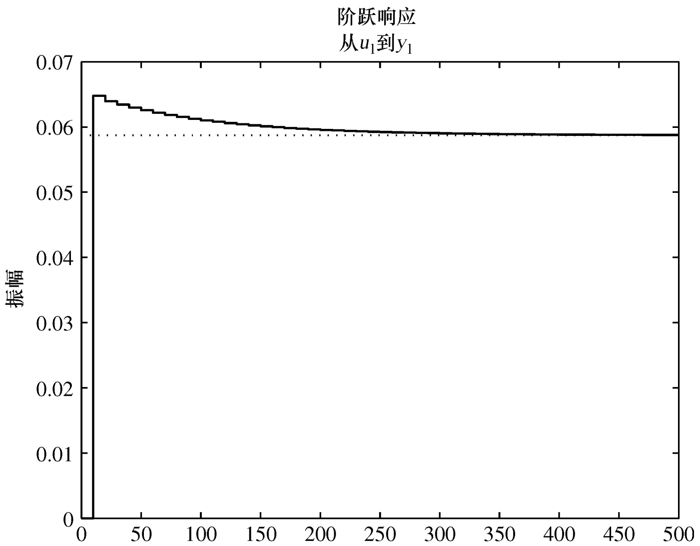


图 9.29 BJ22221 模型阶跃响应（模型 2）

作为比较，所辨识的 BJ 模型为模型 1 和模型 2 在 3 个不同的操作模式分别总结于表 9.6。

表 9.6　3 种不同操作模式的 BJ 模型

	模　型　1	模　型　2
LL	$y_t = \frac{-1.399z^{-1}}{1-0.8979z^{-1}}u_t + \frac{1-0.783z^{-1}}{1-0.9908z^{-1}}e_t$	$y_t = \frac{0.08725z^{-1}-0.07938z^{-2}}{1-0.8956z^{-1}-0.001105z^{-2}}u_t + \frac{1+0.8493z^{-1}+0.01498z^{-2}}{1+0.5083z^{-1}-0.2532z^{-2}}e_t$
NL	$y_t = \frac{-1.341z^{-1}}{1-0.9014z^{-1}}u_t + \frac{1-0.6947z^{-1}}{1-0.8969z^{-1}}e_t$	$y_t = \frac{0.0648z^{-1}-0.05901z^{-2}}{1-0.8976z^{-1}-0.00373z^{-2}}u_t + \frac{1-0.9354z^{-1}+0.05099z^{-2}}{1-1.25z^{-1}+0.3602z^{-2}}e_t$
HL	$y_t = \frac{-1.305z^{-1}}{1-0.9011z^{-1}}u_t + \frac{1-0.7593z^{-1}}{1-0.9676z^{-1}}e_t$	$y_t = \frac{0.04006z^{-1}-0.03878z^{-2}}{1-0.9571z^{-1}-0.008635z^{-2}}u_t + \frac{1+0.9048z^{-1}-0.00262z^{-2}}{1+0.5577z^{-1}-0.3024z^{-2}}e_t$

9.2.6　非线性系统辨识

非线性系统辨识中也试图为燃料流量和电池的温度（模型 1）的过程提供帮助。一个多层次的输入序列是专为非线性系统辨识而定的。输入—输出数据显示在图 9.30。

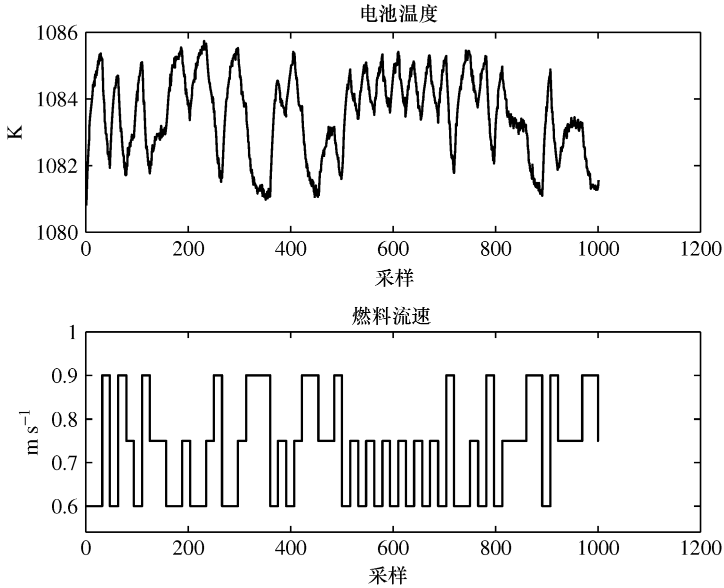


图 9.30　非线性系统辨识的输入—输出数据

作为第一步，开始线性建模。为了便于说明，专注于 BJ 模型。二阶 BJ 模型的交叉验证的结果如图 9.31 和图 9.32 所示。虽然在残差测试中该模型是成功的，并

且百分比也比较高的，图9.31清楚地显示了预测一些一致的不匹配的反应，特别是围绕采样825~850的瞬间。阶数增加1使三阶BJ模型可尝试。交叉验证结果如图9.33和图9.34所示，从而可以得到二阶BJ模型类似的结论。因此，现在看来，增加模型阶数不会造成系统不匹配。

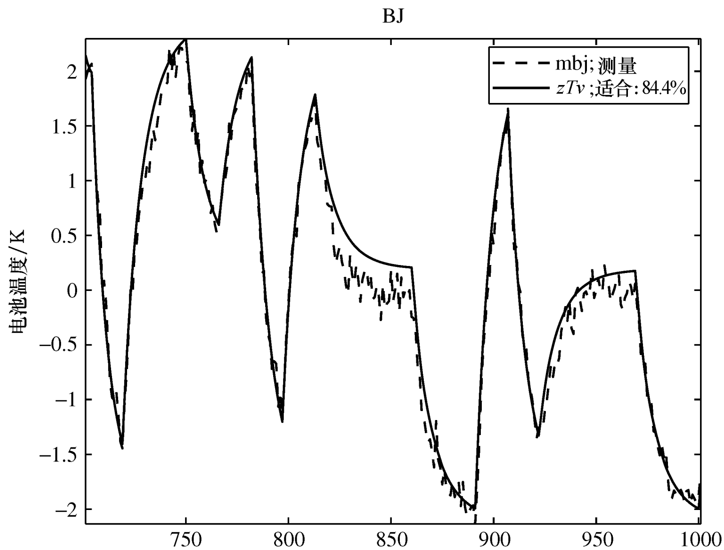


图9.31　BJ22221预测与来自验证数据集的实际电压响应（模型1）

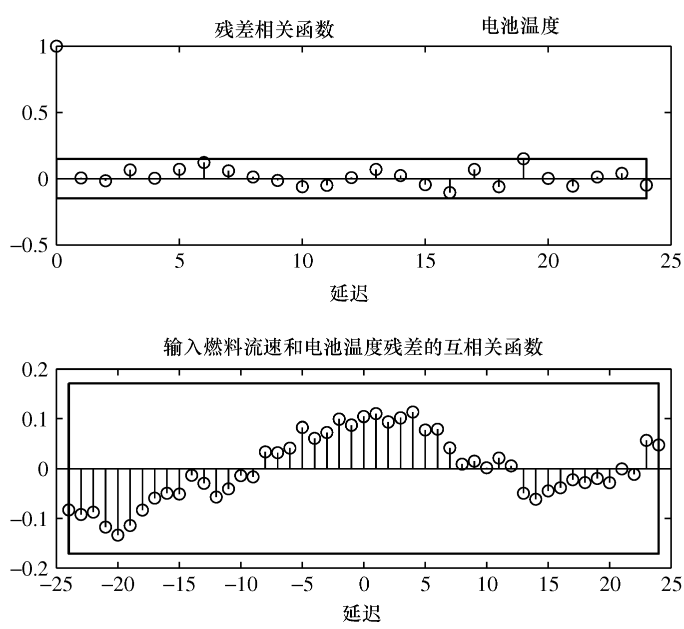


图9.32　基于验证数据集的BJ22221残差测试（模型1）

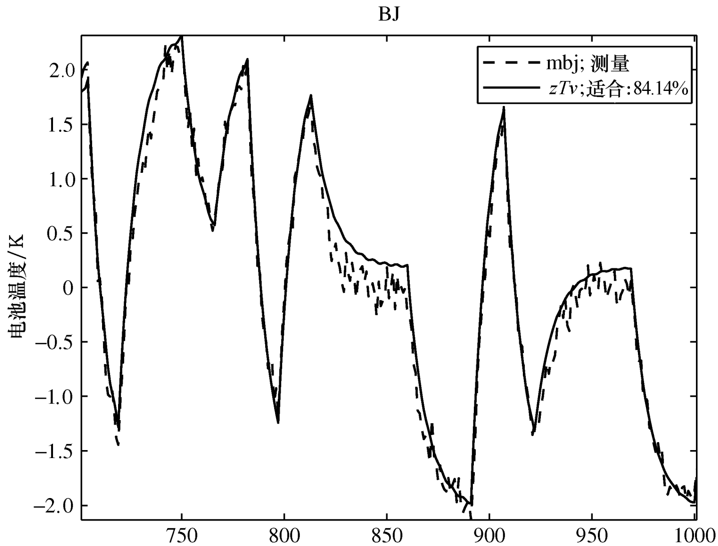


图 9.33　BJ33331 预测与来自验证数据集的实际电压响应（模型 1）

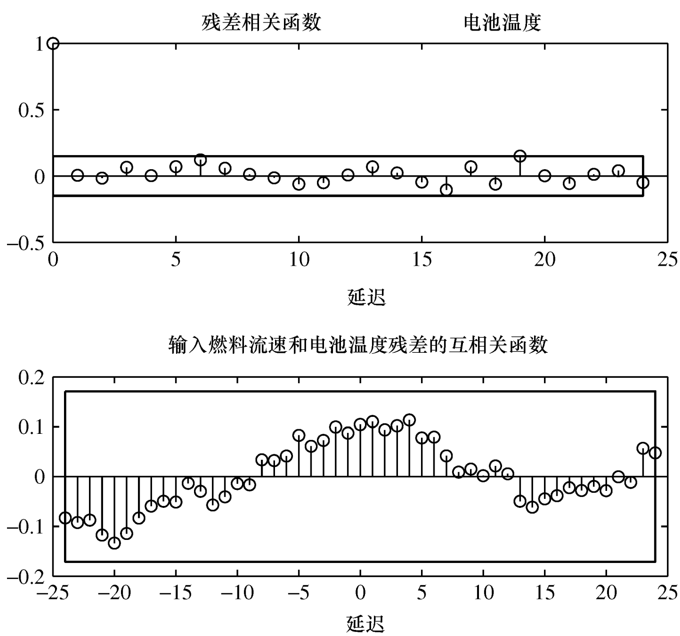


图 9.34　基于验证数据集的 BJ33331 残差测试（模型 1）

三阶非线性 ARX 模型通过以下函数辨识：

mnlarx331 = nlarx（*zTe*，［3 3 1］）

resid（mnlarx331，*zTv*）

compare（mnlarx331，*zTv*）

没有必要去除数据对象 *zTe* 和 *zTv* 做非线性系统辨识。交叉验证的结果示于图 9. 35 和图 9. 36。当与线性辨识结果比较时，可以从图 9. 35 看出，系统的不匹配已经由非线性模型中移除。也可以从图 9. 36 看出，残差测试中的模型是成功的。

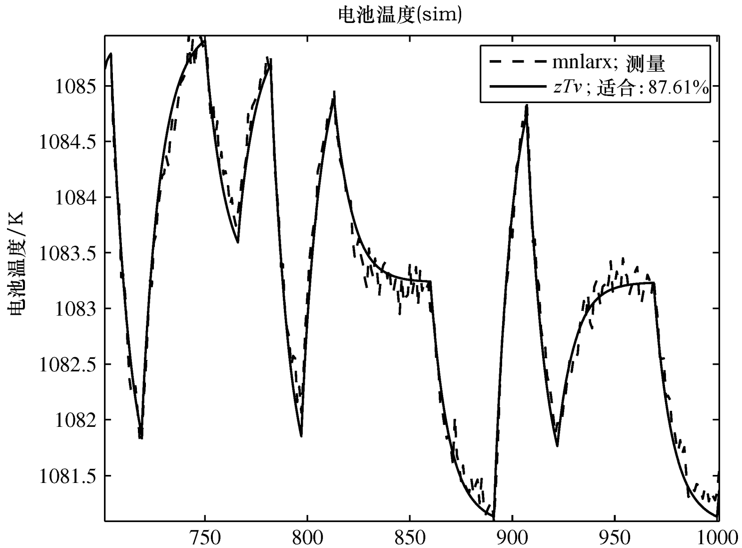


图 9. 35　NLARX331 预测与来自验证数据集的实际电压响应（模型 1）

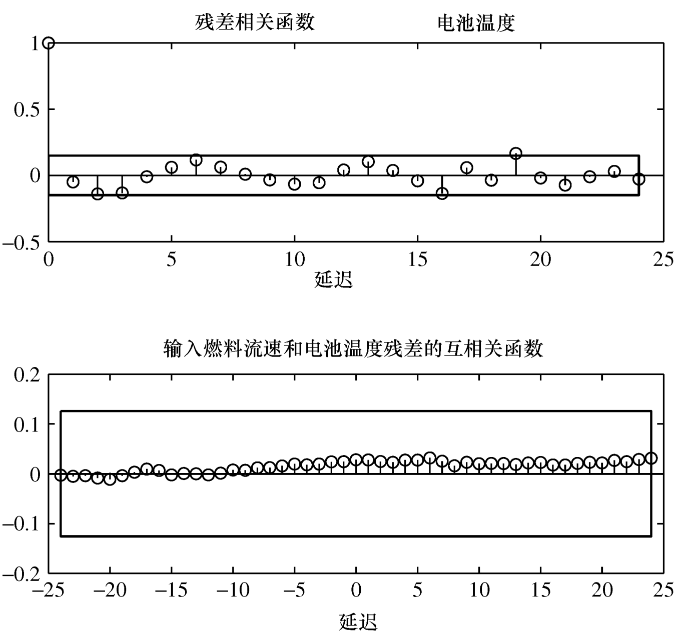


图 9. 36　基于验证数据集的 NLARX331 残差测试（模型 1）

9.3　PID 控制

所辨识的模型 1 是用来设计 PID 控制操作闭环过程的。闭环的结构示于图 9.37 中。该过程的输出被反馈，并与用于输出的设定点进行比较。实际的输出和设定点的误差传送到控制器。该控制器是一个典型的 PI（比例积分）控制器。IMC（内部模型控制）调整方法用于确定基于所述传递函数的控制器，整定参数由系统辨识估计。干扰也加入到这个过程。在这种研究中，过程输出是电池温度。初始稳态值约为 1081 K，设定点再改为 1085 K。

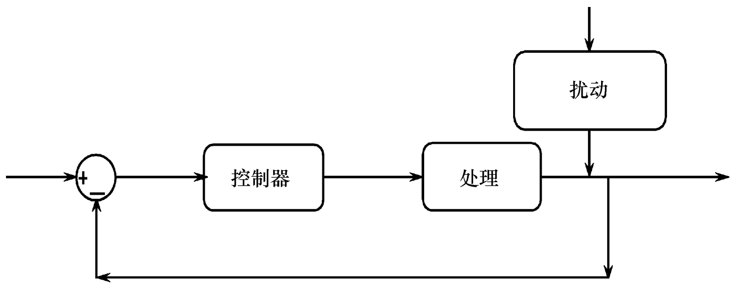


图 9.37　闭环控制框图

对于一阶加纯滞后模型：

$$G(s) = \frac{K}{\tau s + 1}e^{-\tau_d s}$$

IMC 调整方法给出了 PID 控制器参数：

$$K_c = \frac{\tau + \frac{\tau_d}{2}}{K\left(\tau_c + \frac{\tau_d}{2}\right)}$$

$$\tau_I = \tau + \frac{\tau_d}{2}$$

$$\tau_D = \frac{\tau\tau_d}{2\tau + \tau_d}$$

式中，τ_c 是一个用户定义的闭环时间常数。

基本上没有时间延迟（1 个采样延迟是由于零阶保持），因此调整参数成为

$$K_c = \frac{\tau}{K\tau_c}$$

$$\tau_I = \tau$$

$$\tau_D = 0$$

IMC 调整方法是在连续时间域的一阶加纯滞后传递函数。然而，由系统辨识得

到的差分方程是离散时间形式。在这种情况下，由离散时间形式到连续时间形式的转换是必需的。对于离散传递函数：

$$G(z)=\frac{az^{-1}}{1-bz^{-1}}$$

在一个连续的时间域其相应的传递函数可以近似为（假设存在零阶保持）

$$G(s)=\frac{\left(\frac{a}{1-b}\right)}{-\frac{\tau_s}{\ln(b)}s+1}$$

考虑从 OE111 确定的过程模型，它由下式给出：

$$\frac{y_t}{u_t}=\frac{-1.296z^{-1}}{1-0.9071z^{-1}}$$

将其转换为连续时间传递函数得

$$G(s)=\frac{-13.95}{102.56s+1}$$

根据连续的时间传递函数，选择所需的闭环时间常数为 2/3 的开环时间常数，即 $\tau_c=68s$。然后，PID 控制被设计成

$$K_c=-0.1081$$
$$\tau_I=102.56$$
$$\tau_D=0$$

这是一个 PI 控制器，因此 PI 控制可以写成

$$G_c(s)=-0.1081\left(1+\frac{1}{102.56s}\right)$$

9.3.1 设定点跟踪

图 9.38 示出了在额定负载下的闭环响应。设定点的电池温度的变化为 t = 1000s 时 1081 ~ 1085K。整个闭环时间常数大约是 60s，它接近设计的闭环时间常数规范。因此，控制器调整看来是令人满意的，这间接表明了从系统中辨识该模型与实际过程的一个良好近似。

图 9.39 示出了在所有 3 个负载上的闭环响应。闭环时间的低负荷常数约为 95s，用于高负荷约为 43s，而对于额定负载大约是 60s。人们可以看到，PI 控制器的设计在额定负载产生闭环时间常数是最接近所设计的 68s 的闭环时间常数。

9.3.2 抗干扰

同时，进行了抗干扰测试。首先，该过程从设置在 1081K 的相同 PI 控制器开始。图 9.40 显示了不同操作负载的温度分布。作为干扰，在 1000s 时间内燃料气温度从 823K 升到了 853K。正如预期，PI 控制器仿真表明最好的抗干扰性能可在额定负载时实现。这并不令人惊讶，因为这是根据额定负载确定的模型设计出的控制器。

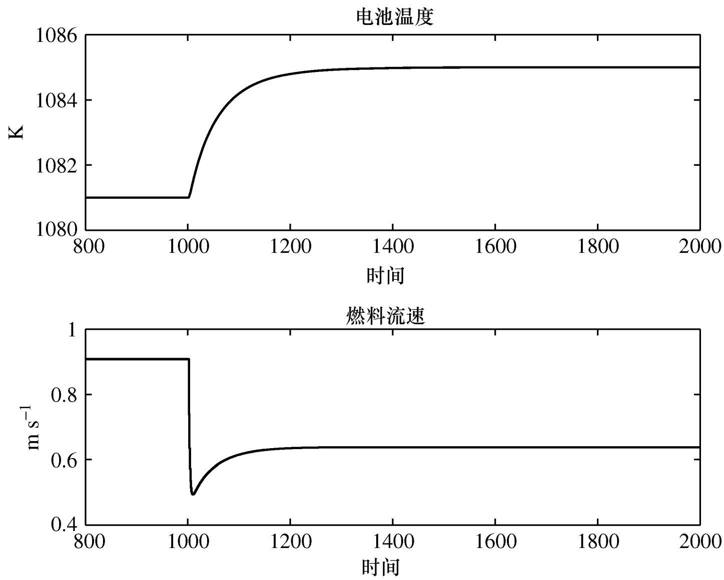


图 9. 38　闭环温度和对在额定负载设置点变化的燃料流量响应

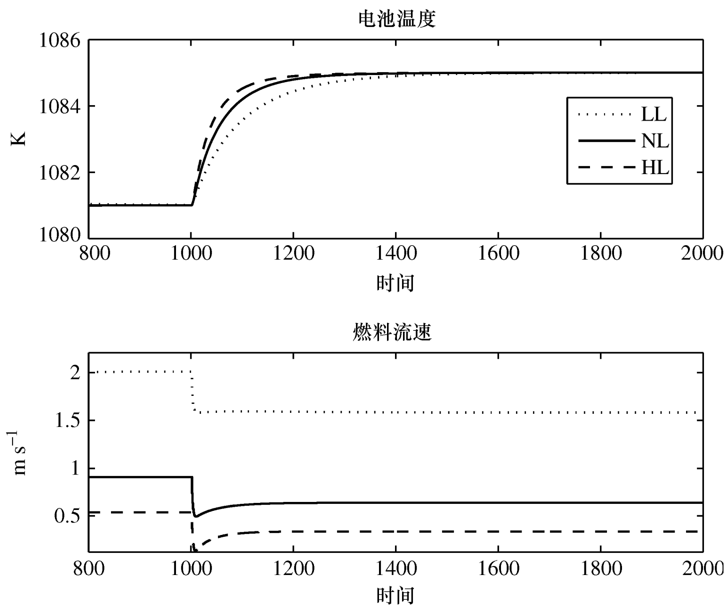


图 9. 39　闭环温度和对在不同负载设置点变化的燃料流量响应

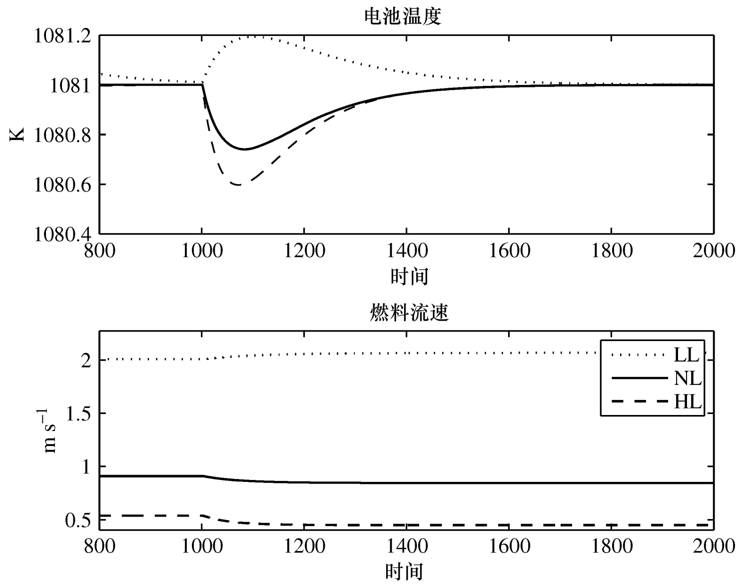


图 9.40　闭环温度和对在不同负载干扰下的燃料流量响应

9.3.3　离散时间过程 IMC

9.3.3.1　完美的控制

在本节中，从辨识的离散时间模型直接考虑控制器的合成。要做到这一点，需要回顾离散时间 IMC。

考虑图 9.41 所示的常规反馈控制系统。

假设设计一个控制器，给出了一个输入信号：

$$u = \frac{1}{G}[r - v] \tag{9.1}$$

输出响应为

$$\begin{aligned} y &= G\left[\frac{1}{G}[r - v]\right] + v \\ &= [r - v] + v \\ &= r \end{aligned}$$

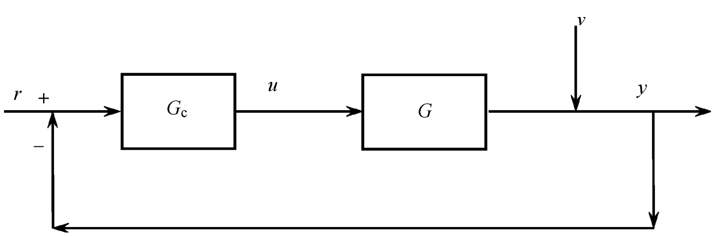


图 9.41　常规反馈控制

这是一个完美的控制，这个完美的控制要求如下：

1）有过程模型的精确知识；

2）过程噪声已测量；

3）过程模型是可逆的。

一般情况下，不能有过程和噪声动态这样的完美的知识。此外，离散时间过程模型永远不能因为时间完全可逆延迟。否则，将有一个非因果项在这样的控制器传递函数中，这违反了因果关系（不是物理可实现的）。尽管存在这些的问题，但已经看到了如何结合现有的工艺知识（以模型的形式）到控制器可以提高性能。

妨碍实现完美控制的一些问题可以归纳如下：

1）不可测干扰；

2）时间延迟；

3）不稳定零点；

4）植物模型不匹配；

5）过多的控制动作。

在本节中，将讨论如何通过引入控制来处理所有这些问题设计框架被称为 IMC。一旦这 5 个问题都解决了，IMC 控制器的设计也完成了。

9.3.3.2　基本 IMC

在许多基于模型的控制方法中，该过程模型被用于从给定的操作变量的值（u）预测输出（y）。传统的反馈控制器配置可以稍微进行修改以适应过程模型，G_m 如图 9.42 所示。

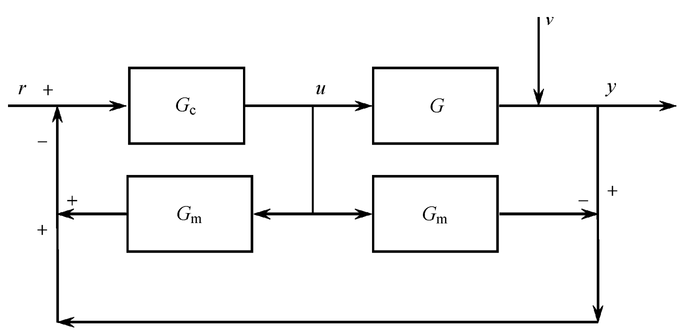


图 9.42　基本 IMC

这是完全等同于标准的反馈配置。大体上相加和相减过程模型的输出，然后可以将控制动作分成几个分量的动作，如图 9.43 所示。

图 9.43 所示的结构可进一步简化成基本 IMC 形式，如图 9.44 所示，其中

$$C = \frac{G_c}{1 + G_c G_m} \tag{9.2}$$

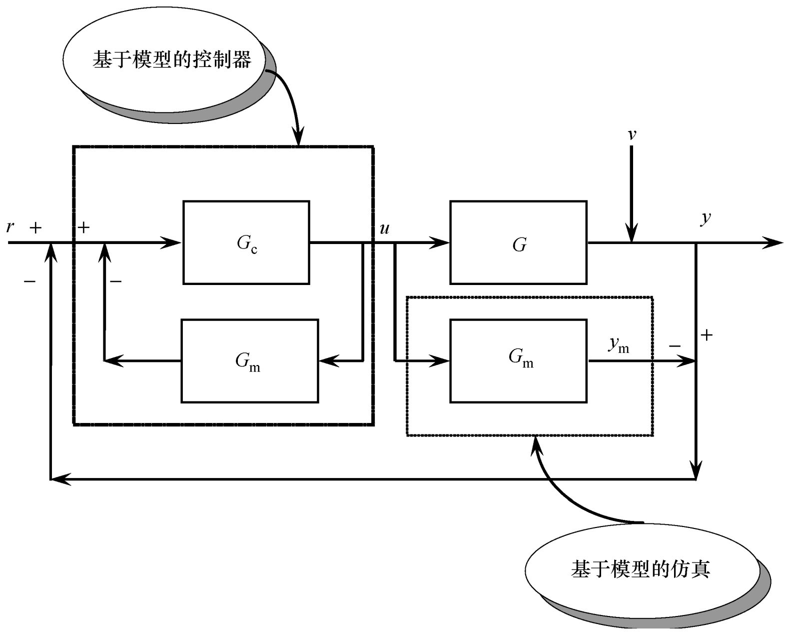


图 9.43　分离成几个组件动作的控制作用

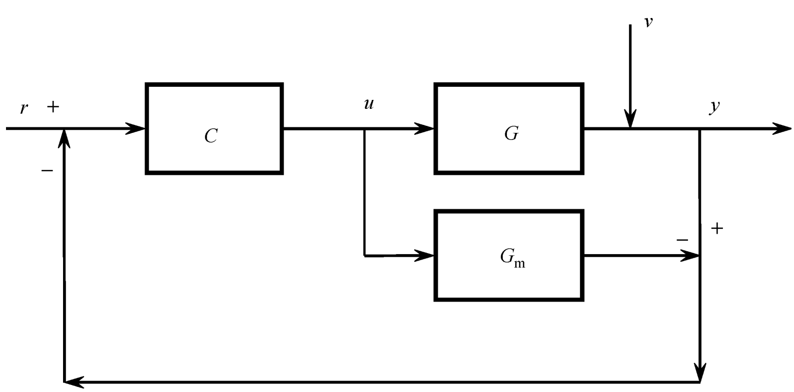


图 9.44　基本 IMC 组件

上述 IMC 控制等效于传统的反馈控制，如图 9.45 所示，其中

$$G_c = \frac{C}{1 - CG_m} \tag{9.3}$$

请注意，在 IMC 中，而不是反馈的输出（y）之间的差值测量和预测输出（$y - y_m$）反馈。然后考虑当模型精确时，也就是是在 $G_m = G$ 时，反馈仅仅关于过程噪声（v），如图 9.46 所示。

在图 9.46 的基础上，响应可以很容易地计算为

$$
\begin{aligned}
y &= GC[r-v]+v \\
&= GCr+(1-GC)v
\end{aligned}
$$

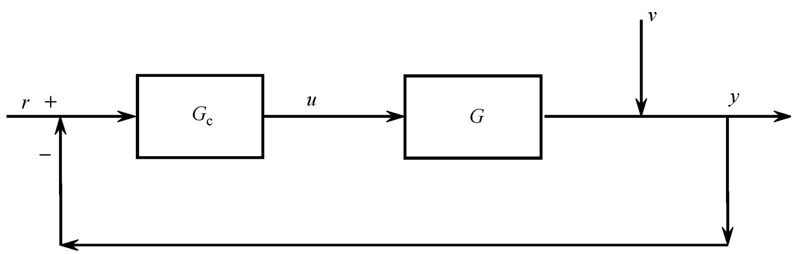


图 9.45　常规反馈控制 IMC 等价

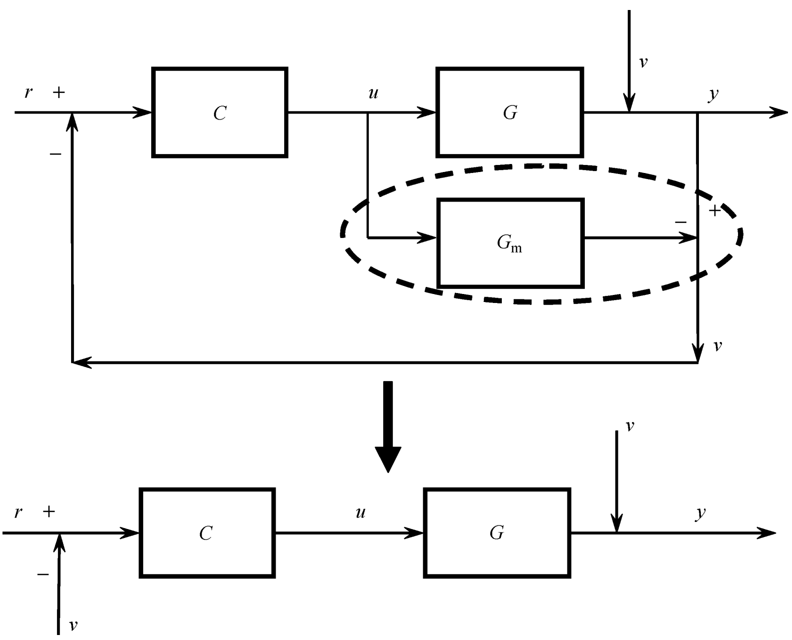


图 9.46　设计完美控制 IMC 变换

利用这一办法，在“完美”的控制设计中尝试其中一个重要的问题，即不可测量的扰动已根据 IMC 框架“测量”间接的影响。

现在可以设计一个完美的控制器：

$$C=\frac{1}{G_m} \tag{9.4}$$

得到

$$y=r+0v=r$$

然后将该 IMC 控制器传送回 G_c。但是，如果有时的时间延迟了 G_m，则控制器不是按规律实现的。在这种情况下，可以分解模型为两部分：G_{m_-} 不包含的时间延迟和 G_{m_+} 延迟部分。例如，如果一个过程模型可以写成

$$G_m = \frac{z^{-2}(1-0.5z^{-1})}{1-0.9z^{-1}}$$

则有

$$G_{m_+} = z^{-2}$$

$$G_{m_-} = \frac{1-0.5z^{-1}}{1-0.9z^{-1}}$$

作为一个结果，IMC 控制器可设计成

$$C = \frac{1}{G_{m_-}} = \frac{1-0.9z^{-1}}{1-0.5z^{-1}}$$

因此，在一般情况下，如果该方法具有时间延迟，则首先需要分解出这样的延迟：

$$G_m = z^{-d} G_{m_-}$$

式中，G_{m_-} 不包含任何时间延迟。

然后 IMC 控制器可设计成

$$C = \frac{1}{G_{m_-}}$$

输出响应现在由下式给出：

$$y_t = z^{-d} r_t + (1 - z^{-d}) v_t = r_{t-d} + (v_t - v_{t-d})$$

它被称为有限拍控制，这是在时间延迟的存在下可能的最佳控制。

正如前面讨论的那样，并不是所有的模型动态是可逆的。模型可以含有非最小相位零点（非可逆零点）。作为控制器，C 必须是稳定和可实现的，该过程模型将估算为非可逆和可逆部分的结果：

$$G_m = G_{m_+} G_{m_-} \tag{9.5}$$

非可逆的动态和时间延迟相结合，形成了 G_{m_+}，而使所有在 G_{m_-} 可逆动态。

q 个非最小相位零点加上延迟的因式分解由 Garcia 和 Morari（1982）提出：

$$G_{m_+} = z^{-d} \prod_{i=1}^{q} \left[\frac{1-\xi_i z^{-1}}{1-\dfrac{1}{\xi_i} z^{-1}} \times \frac{1-\dfrac{1}{\xi_i}}{1-\xi_i} \right]$$

式中，ξ_i 为位于单位圆在 z 以外的非最小相位零点。

这等同于用它们在单位圆内的映像来替换所有的非最小相位零点。

然后，控制器被定义为

$$C = \frac{1}{G_{m_-}}$$

最后，为了保证统一的闭环增益，稳态控制增益必须是稳态模型增益的倒数，因此需要检查：

$$C(1)=\frac{1}{G_{\mathrm{m}}(1)}$$

综上所述，基本 IMC 设计包括以下内容：

1）模型为可逆和不可逆部分；

2）定义控制器作为模型的一部分可逆的逆；

3）检查稳态控制增益。

9.3.3.3　抗振铃基本 IMC

每当一个控制器有一个极点接近值 -1，控制器振铃可发生。例如，考虑一个控制器传递函数为

$$u_t=\frac{1}{1+0.9z^{-1}}e_t$$

有一个极点 -0.9 接近 1。假设 e_t 只是一个脉冲，即 $e_t=\delta_t$，则 y_t 的响应可用长除法计算

$$\frac{u_t}{\delta_t}=\frac{1}{1+0.9z^{-1}}=1-0.9z^{-1}+0.81z^{-2}-0.729z^{-3}+\cdots+$$

即

$$\begin{cases}u_0=1\\u_1=-0.9\\u_2=0.81\\u_3=-0.729\\\vdots\end{cases}$$

请注意，该控制器的输出在每个采样间隔改变它的轨迹。这类控制器输出被视为控制振铃，控制器振铃会严重磨损出阀门，应当在设计阶段避免。

这是必要的，以避免在设计中考虑到 IMC 控制器铃声控制器。解决的办法很简单：不反转是接近 -1 的过程零。例如，考虑一个过程传递函数为

$$G_{\mathrm{m}}=\frac{z+0.9}{z^2-0.8z}$$

这个传递函数可以写成

$$G_{\mathrm{m}}=\frac{z^{-1}(1+0.9z^{-1})}{1-0.8z^{-1}}$$

有一个时间延迟，但没有不可逆的零点。然而，传递的零功能，-0.9，接近 -1，不应该被反转。因此，不得不将时间延迟，连同“振铃”因子（$1+0.9z^{-1}$）相结合，形成不可逆的该过程传递函数（G_{M_+}）的一部分，也就是

$$G_{m_+}=z^{-1}(1+0.9z^{-1})\frac{1}{1+0.9}$$

因此

$$G_{m_-} = G_m / G_{m_+} = \frac{1.9}{1 - 0.8z^{-1}}$$

并且（抗振铃）基本 IMC 控制器为

$$C = 1/G_{m_-} = \frac{1 - 0.8z^{-1}}{1.9}$$

检查是否

$$\lim_{Z^{-1}\to 1} CG_m = 1$$

值得重申的是，非可逆零的因式分解与零点之间的不同，这一值是接近 -1 的。为了说明这一点，考虑一个例子：

$$G_m = \frac{z^{-1}(1 - 1.25z^{-1})(1 + 0.9z^{-1})}{1 - 0.9z^{-1}}$$

这其中有一个时间延迟，一个非可逆零是接近 -1 的。G_{m_+} 应该是

$$G_{m_+} = z^{-1}(1 + 0.9z^{-1})\frac{1}{(1 + 0.9)} \times \frac{(1 - 0.25z^{-1})}{(1 - (1/1.25)z^{-1})} \times \frac{(1 - 1/1.25)}{(1 - 1.25)}$$

因此，可以计算出过程传递函数的可逆部分，然后设计的基本 IMC 控制器是稳定的、物理可实现的和不振铃的。

9.3.3.4　强大的 IMC 设计

IMC 的主要优点是反馈信号中包含明确的关于植物模型不匹配的信息，这使 IMC 直接解决“鲁棒性”问题。

已经看到，在一些模型反演控制器（例如无差拍），所需要的控制作用可以是积极的。此外，使用基本 IMC 设计的产品，植物模型不匹配可能会导致控制器的不稳定。Garcia 和 Morari（1982，1985）建议在控制回路中引入过滤器，如图 9.47 所示。

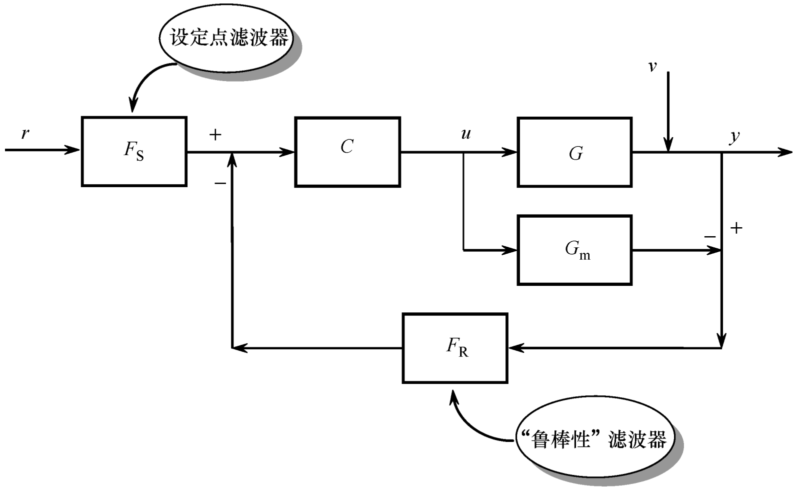


图 9.47　IMC 滤波器

设定点滤波器 F_S 是选择提供一个参考轨，这与工艺操作的目标一致。它经常被选择为

1）$F_S = 1$（使设定点通过）

2）$F_S = \dfrac{1-\lambda}{1-\lambda z^{-1}}$（低通过过滤设定点）

这是增益 =1 的一阶传递函数。$\lambda(0<\lambda<1)$ 可以从采样时间和时间常数定义如下：

$$\lambda = \exp(-T_S/\tau)$$

式中，T_S 是采样时间；τ 是滤波器的时间常数。

Morari 和 Zafiriou（1989）详细讨论了“鲁棒性”过滤器 F_R 的设计。然而，在一般情况下，对于具有至多一个积分项（即至多有一个极点位于 $z=1$）的过程中，推荐的形式是低通滤波器：

$$F_R = \frac{1-\alpha}{1-\alpha z^{-1}} \quad 0<\alpha<1$$

被选择的滤波器常数 α 以确保闭环系统在稳定预期范围植物模型不匹配。注意，当 α 趋近于零时，信号被反馈不用过滤；而当 α 趋近于 1 时，很少有反馈。然后，滤波器的设计是性能和鲁棒性之间的折衷。

在简单的情况下，$G=G_m$，过滤的 IMC 控制可被简化，如图 9.48 所示。

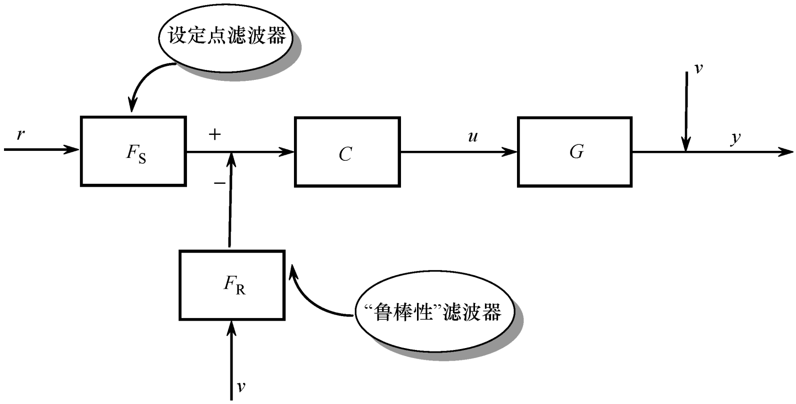


图 9.48　简化 IMC 滤波器

闭环过程响应可表示为

$$y = F_S C G_m r + (1 - F_R C G_m) v$$

当且仅当传输函数 G_m、C、F_S 和 F_R 稳定时，该闭环系统是稳定的。

闭环动态可以通过调整设置点滤波器和“鲁棒性”滤波器进行调谐。事实上，这两个过滤器发挥在这种情况下相同的作用，这是为了减小的响应速度和控制动作的幅度。

例如，闭环响应于设定点的改变取决于设定点过滤器。如果希望闭环响应为 τ 的时间常数，则该设定点滤波器可以通过如下计算：

$$F_S = \frac{1-\lambda}{1-\lambda z^{-1}}$$

式中
$$\lambda = \exp(-T_S/\tau)$$

一般来说，较大的 λ，则较慢的响应，同时，较小的控制动作。

在大多数应用中，“鲁棒性”滤波器被选择为相同的设定点滤波器。对于这种情况下的框图变得进一步简化，如图 9.49 所示。

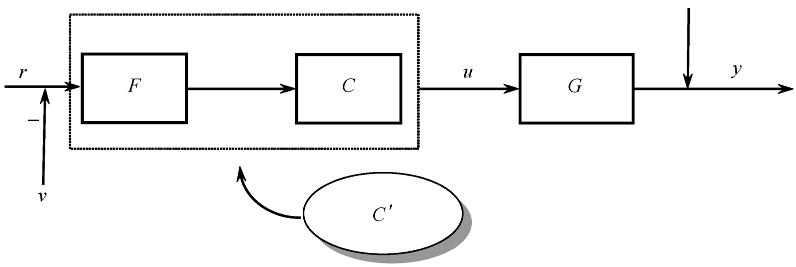


图 9.49　进一步简化 IMC 滤波器

那么这个新的 C' 被定义为一个鲁棒 IMC 控制器。它可以被传递到传统的反馈控制的关系中：

$$G_c = \frac{C'}{1-C'G_m}$$

9.3.4　多环控制 SOFC 离散时间 IMC 应用

在本节中，设计了多回路的离散时间 SOFC 模型验证前面的章节。考虑表 9.6（中排）所示的标识额定负载 BJ 模型。与燃料流量相关的电池温度的模型由下式给出：

$$\frac{y_t}{u_t} = G_{p1}(z^{-1}) = \frac{-1.341z^{-1}}{1-0.9014z^{-1}}$$

与燃料流量的压力相关的电压模型由下式给出：

$$\frac{y_t}{u_t} = G_{p2}(z^{-1}) = \frac{0.0648z^{-1}-0.05901z^{-2}}{1-0.8976z^{-1}-0.00373z^{-2}}$$

这两种模式的无延迟部分可以写成

$$G_{p1_}(z^{-1}) = \frac{-1.341}{1-0.9014z^{-1}}$$

$$G_{p2_}(z^{-1}) = \frac{0.0648-0.05901z^{-1}}{1-0.8976z^{-1}-0.00373z^{-2}}$$

再考虑到所需的闭环时间常数 $\tau_c = 68\text{s}$。与采样时间 $T_s = 10\text{s}$，“鲁棒性”过滤器对于 IMC 可设计成

$$F_{\mathrm{R}}=\frac{1-\alpha}{1-\alpha z^{-1}}$$

其中，基于期望闭环时间常数和采样时间，滤波器参数 α 可以被确定为

$$\alpha=\exp(-T_{\mathrm{s}}/\tau_{\mathrm{c}})=\exp(-10/68)=0.8623$$

从而得到

$$F_{\mathrm{R}}=\frac{0.1368}{1-0.8623z^{-1}}$$

离散时间 IMC 因此可以被确定为

$$C_1'=F_{\mathrm{R}}\frac{1}{G_{p1-}(z^{-1})}=\frac{-0.1020+0.0919z^{-1}}{1-0.8623z^{-1}}$$

$$C_2'=F_{\mathrm{R}}\frac{1}{G_{p2-}(z^{-1})}=\frac{0.1368-0.1228z^{-1}-0.00051z^{-2}}{0.0648-0.1149z^{-1}+0.05088z^{-2}}$$

最后，将离散时间 IMC 表示为传统的反馈控制可以确定

$$G_{\mathrm{c}}=\frac{C'}{1-C'G_{\mathrm{p}}}$$

例如，对于模型 1 传统的反馈控制：

$$G_{\mathrm{c1}}(z^{-1})=\frac{-0.102+0.1838z^{-1}-0.08284z^{-2}}{1-1.9013z^{-1}+0.9013z^{-2}}$$

为了测试控制性能，首先，这两个 IMC 要单独实施。对于一对燃料流率输入和温度输出，初始标称设定点 1081K 被设置为启动程序。同时在 $t=1000\mathrm{s}$ 时，设置点由 1081K 变为 1083K。图 9.50 所示为 3 个不同的负载闭合响应，这与之前较早的研究连续时间型 PI 控制性能非常相似。对于一对燃料流量压力的输入输出电压，0.83V 的初始标称设置点设置为程序启动，并在 $t=500\mathrm{s}$，设定点从 0.83 变化到 0.84，图 9.51 所示是额定负载闭环响应。其中所述设定点被正确跟踪，闭环稳定时间也满足所需 $\tau_{\mathrm{c}}\approx 68\mathrm{s}$。尽管事实是基于不完善模型所设计的控制器，所辨别的模型不能捕获快速初始逆响应如图 9.3 所示，然而却满足设计规范。然而，不受控制的初始逆响应行为仍存在于闭环响应，如图 9.51 所示。放大的视图显示在图 9.52中，其中，可以清楚地看到逆反应。

最后，这两个 IMC 同时实现测试多回路闭环反应。对电池电压的设定点变成在 $t=1100\mathrm{s}$ 0.83 ~0.84V，而对于电池温度的设定点变成在 $t=800\mathrm{s}$ 1081 ~1083K。多回路闭环响应示于图 9.53 中，其中图 9.53a 和 c 示出了电池的电压响应及其相应的控制器响应（燃料流压力），而图 9.53b 和 d 示出了电池的温度响应及其相应的控制器响应（燃料流率）。由于相互作用，电池电压的峰值与混乱度可以清楚地看出在 $t=800\mathrm{s}$ 时存在于电池温度的设定点的变化。

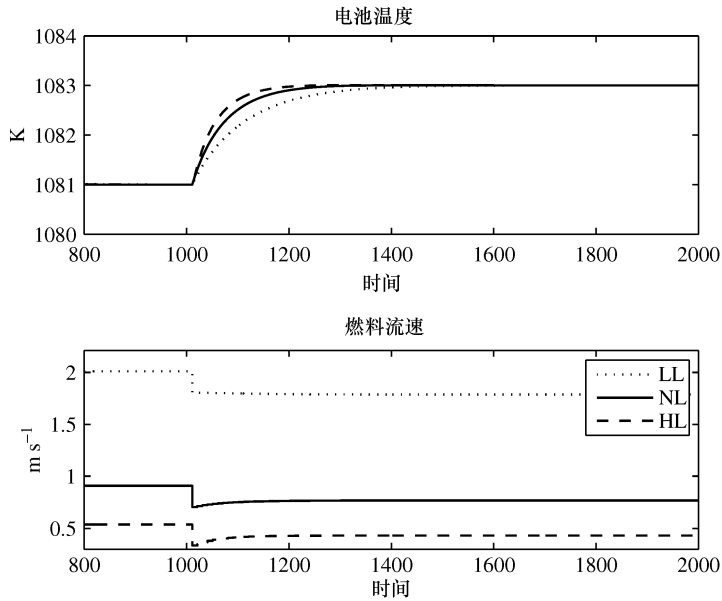


图 9.50 闭环温度和对不同负荷设置点变化的燃料流量响应（离散 IMC 控制）

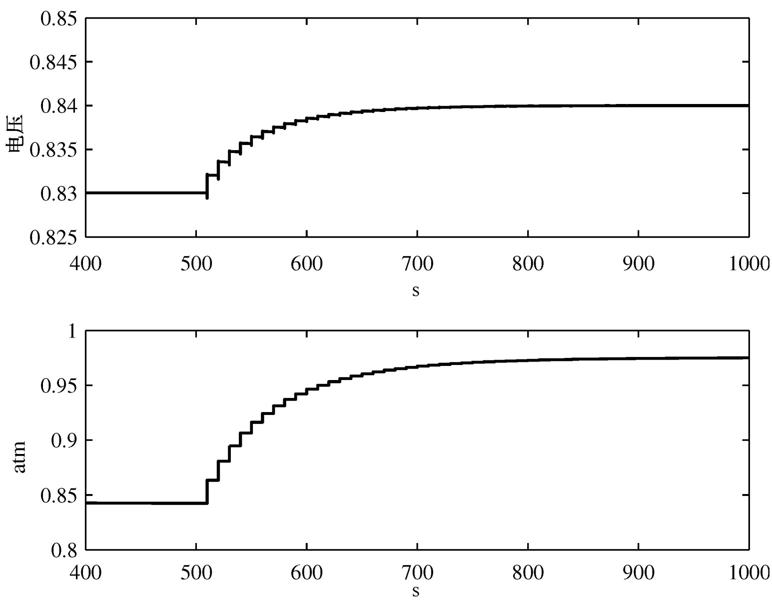


图 9.51 闭环电压和对设定点变化的燃料流量压力响应（离散 IMC 控制）

类似地，在 $t = 1100\mathrm{s}$ 时，电池温度也扰乱了电池电压在设定点的变化。总的来说，多回路闭环响应是令人满意的。这是由于可能在双方控制量小的设定点的改变，使得该过程可以用线性模型来近似。

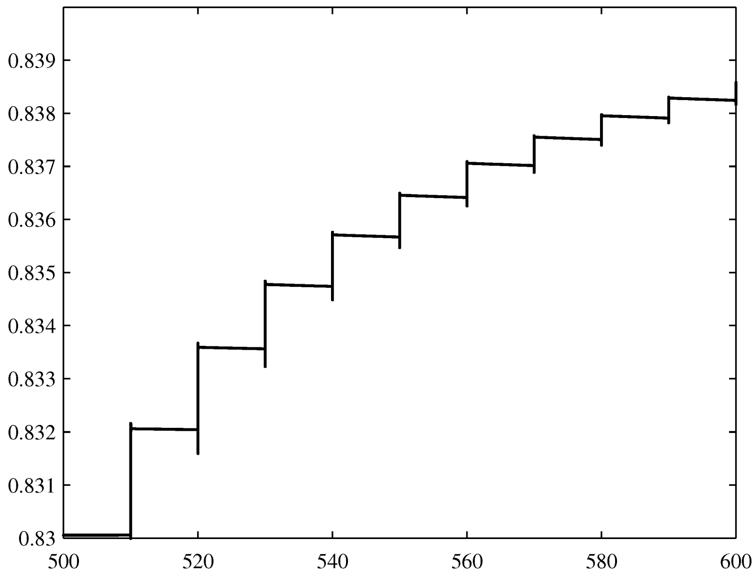


图 9.52　闭环电压的放大视图和对设定点变化的燃料流量压力响应（离散 IMC 控制）

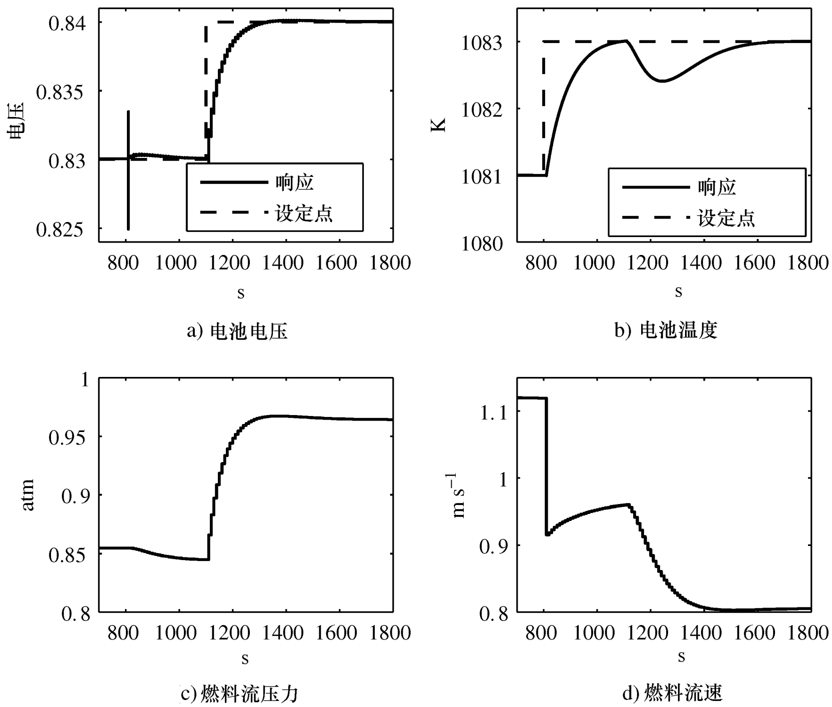


图 9.53　设定点变化的闭环响应（离散多环 IMC 控制）

9.4 闭环辨识

在本节中，当离散时间温度循环 IMC 很到位，以及模型 2 的闭环辨识当离散时间电压回路 IMC 到位时，考虑模型 1 的闭环辨识。由于两个 IMC 控制器产量大致相同的所需的闭环时间常数，约为 68s，输入信号的带宽可被设计成

$$上限 = \frac{2 \times 10\text{s}}{60\text{s} \times \pi} = 0.1061$$

首先考虑离散时间 IMC 温度循环。设计的输入和输出数据对应的外部激励如图 9.54 所示。与燃料流量相关的电池温度的 BJ11111 模型是通过直接闭环辨识方法辨别的。预测和验证数据集的实际响应示于图 9.55 中。残差的测试显示在图 9.56中，从中可以看出，所辨别的模型中的自相关检测失败。即使适合对验证数据集的百分比是 86.54%，如图 9.55 所示，但这种模式可能仍然是值得商榷的，因为它没有通过残差检验。进一步调查是必要的。

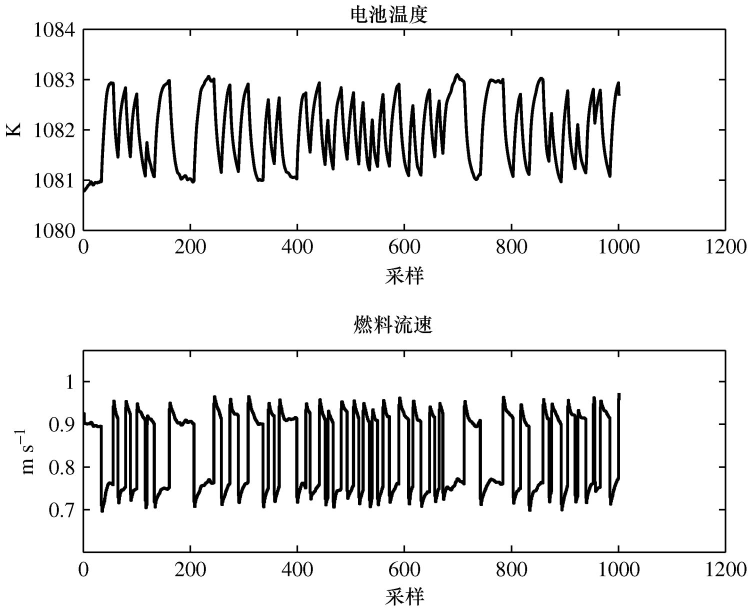


图 9.54 闭环辨识的温度响应数据

BJ11111 在过程和扰动上具有一阶模型结构。既然它未能通过自相关测试，下一步将是将干扰模型的阶数增加 1，使 BJ12211 模型成为下一个候选。从燃料流量的 BJ 模型到电池温度是通过直接闭环辨识的：

$$y_t = \frac{-1.492z^{-1}}{1-0.8926z^{-1}}u_t + \frac{1+0.6638z^{-1}+0.07625z^{-2}}{1-1.594z^{-1}+0.63z^{-2}}e_t$$

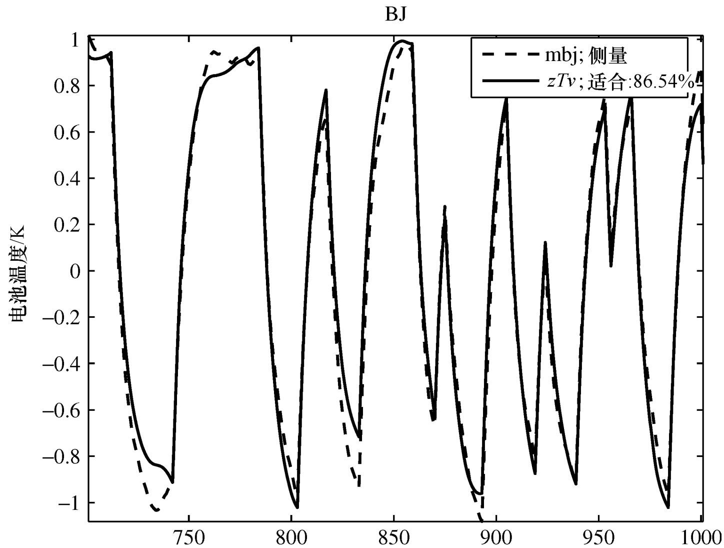


图 9.55　BJ11111 预测与对验证数据集的闭环温度响应

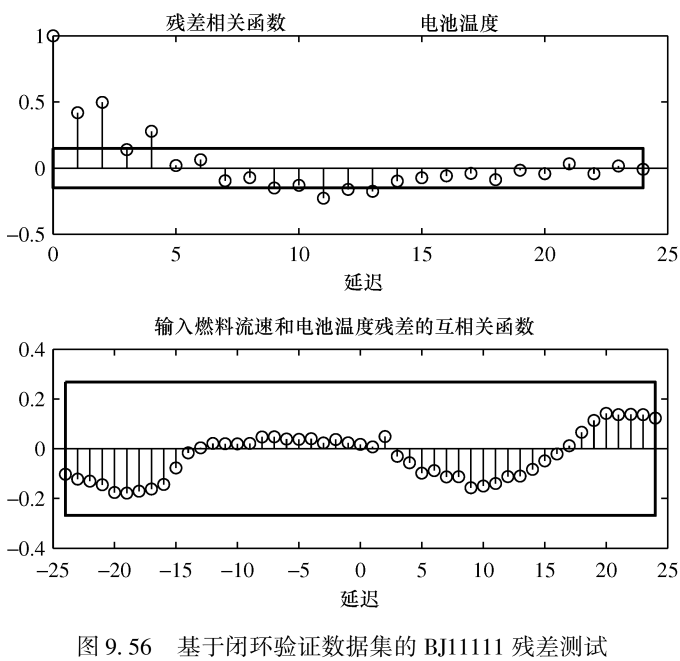


图 9.56　基于闭环验证数据集的 BJ11111 残差测试

该过程模型部分有相似的表9.6 对应（NL 行与模型1 列）。预测的响应和对验证数据集的实际响应示于图 9.57 中。图 9.58 显示了这个模型残差测试成功。

现在考虑离散时间 IMC 电压回路。输入和输出数据设计的外部激励示于图 9.59中。从 BJ11111 模型的燃料流量到电池电压通过直接闭环辨识标识。该残

差测试都显示在图 9. 60 中，可以得出结论，该模型中的互相关测试失败。

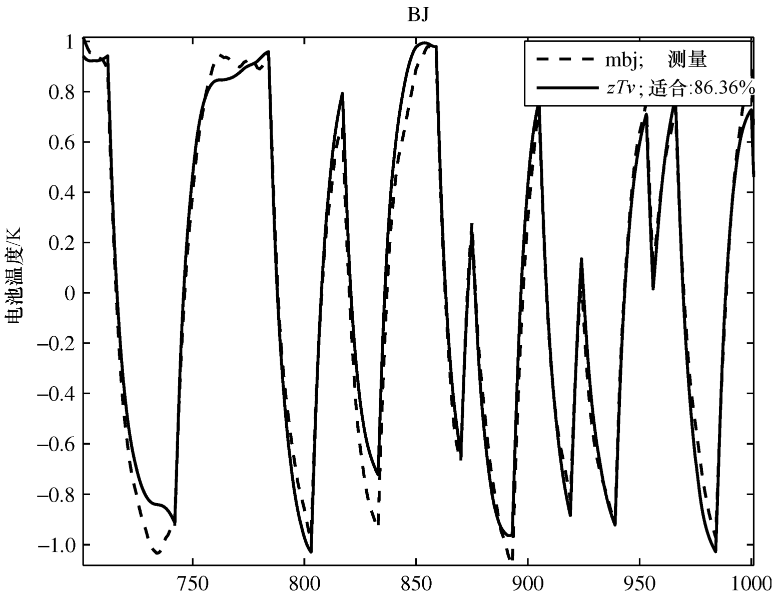


图 9. 57　BJ12211 预测与验证数据集闭环温度响应

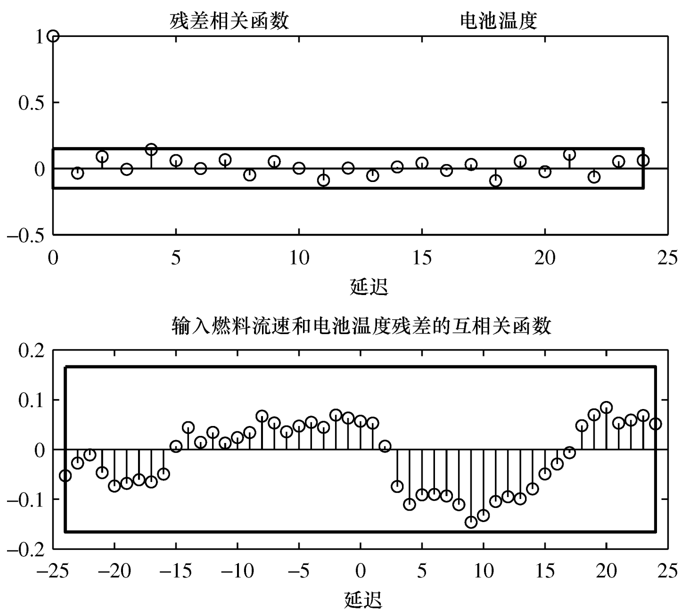


图 9. 58　基于闭环验证数据集的 BJ12211 残差测试

现在，过程和干扰模型均增加一阶，使 BJ22221 模型成为下一个候选。从燃料流的压力到电池电压，BJ 模型被辨识为

$$y_t = \frac{0.08544z^{-1} - 0.07905z^{-2}}{1 - 0.8784z^{-1} - 0.03649z^{-2}}u_t + \frac{1 + 0.8246z^{-1} - 0.03466z^{-2}}{1 + 0.3451z^{-1} - 0.4343z^{-2}}$$

预测并获得在验证数据集的实际响应如图 9.61 所示。图 9.62 显示了这个模型残差测试成功。该过程模型部分有相似的表 9.6 对应（NL 行和模型 2 列）。

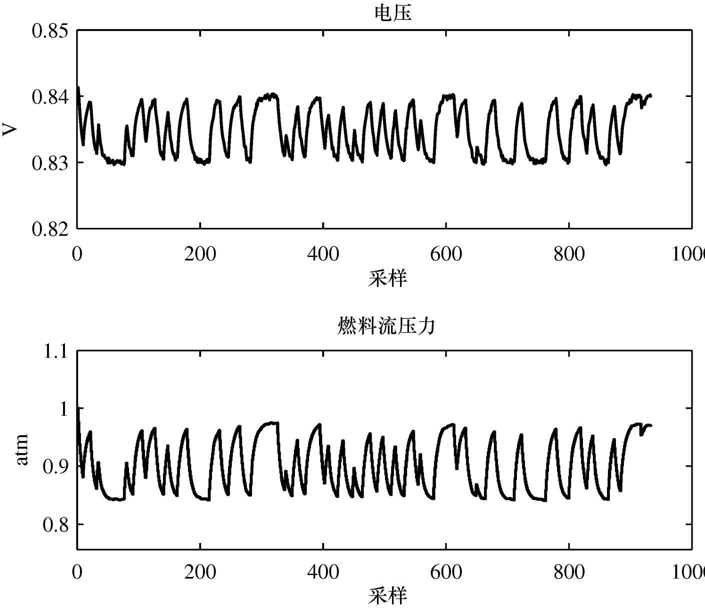


图 9.59　闭环辨识电压响应数据

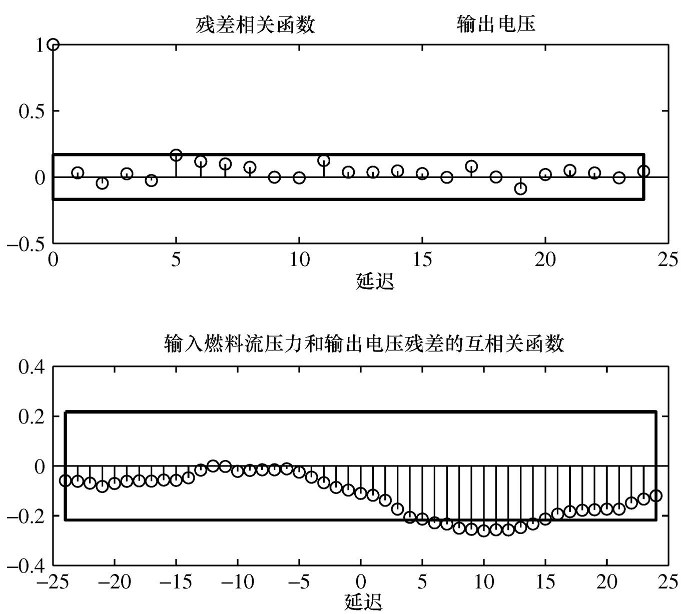


图 9.60　闭环验证数据集 BJ11111 残差测试

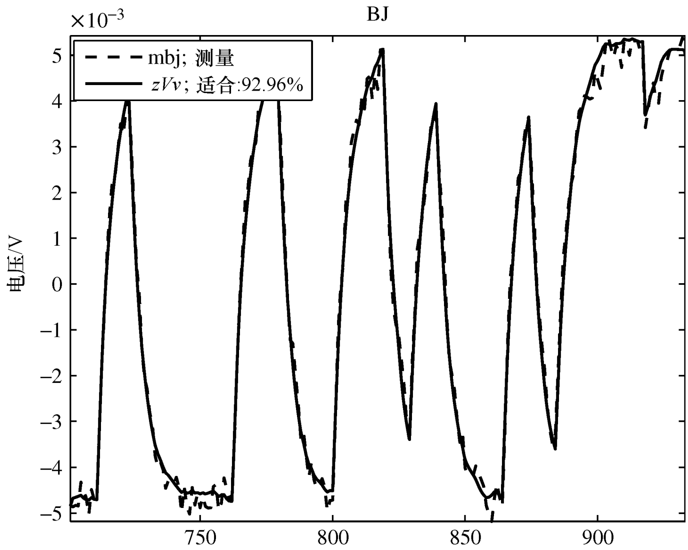


图 9.61　BJ22221 预测与对验证数据集的闭环电压响应

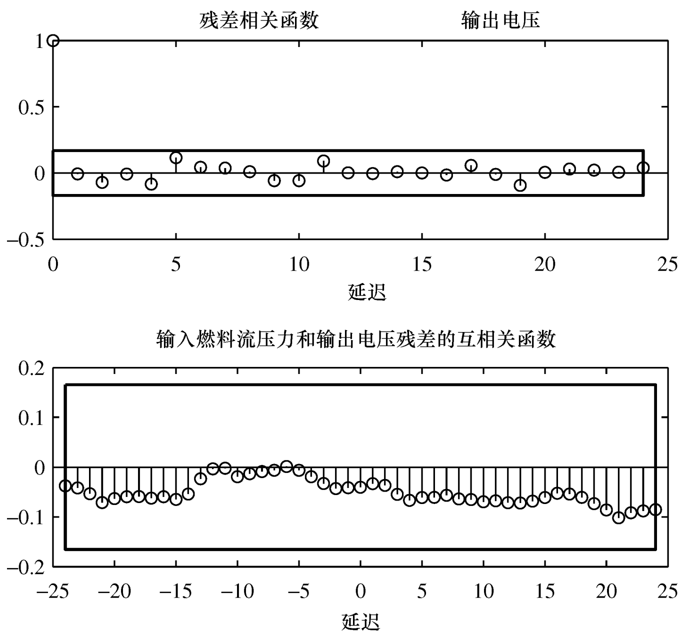


图 9.62　闭环验证数据集 BJ22221 残差测试

9.5　注解和参考文献

系统辨识理论在第 3 章和第 4 章中已有介绍。本章详细示出了 SOFC 建模辨识的应用。输入选择、输入信号设计和初步的步测试，均基于内部技术报告（Wang 和 Huang，2011）。对于离散时间过程内模控制的理论部分是基于阿尔伯塔大学任教时的讲义，“中间过程控制”（Huang 和 Forbes，1997- 2008）。

第 3 部分　板式 SOFC

第 10 章　板式 SOFC 动态建模：第一性原理方法

10.1　简介

在第 7 章已经介绍了 SOFC 的第一性原理建模。管式 SOFC 以超越其他电池的优势而闻名，板式由于高温应用下优异的结构完整性而闻名。然而，同时，它的几何结构在堆叠设计方面提出了挑战，因为相比板式设计，它固有高电阻。在另一个方面，SOFC 电池板很容易构造，而且越来越容易吸引商业关注。

在本章，板式 SOFC 的动态模型，是在第一性原理建模技术的基础上发展起来的。由于模型的复杂性取决于其预定用途，它的范围可以从代数方程组呈现出的简单的线性稳态模型到偏微分方程呈现出的复杂的 3D 动态模型。虽然第一种类型的模型相对于不同的设计之间的性能比较来说可用于模拟稳态响应，但它不能被用于模拟的瞬态响应，或设计一个控制器。另一方面，后一种类型，可用于模拟在空间方向上的瞬态响应，从而可以在任何给定的时间提供深入洞察电池的内部的工艺条件。然而，开发和维护一个复杂的模型需要大量的时间和劳动。基于这种复杂的模型的控制器设计是很复杂的，因此，在多数情况下，这个模型被减少到一个简单的低阶动态模型。由于主要目标是开发一个动态模型，这个模型可以相对容易地设计一个基于模型的控制器，还应该包括以下特点：

1）该模型应该包括一组线性或非线性常微分方程，它可以预测燃料电池的瞬态行为；

2）该模型在基于模型的控制器的设计上应该足够简单，如一个线性或非线性模型预测控制器；

3）它应该能够预测所有重要的工艺条件，包括终端电压、压力和温度；

4）模型对于广泛运营区域应该是有效的，这是建立第一性原理模型的主要驱动因素。

在这些建模目标的基础上，下面的部分将要列出两种类型。首先，一个简单的模型是由假设电池整体温度分布均匀（包括固相和气相）推导出来的。这个模型将被命名为“集中模式”。虽然这种假设在较小的电流，负载电极、互连器和未反

应的气体的温度不会有太大差别的情况下是相当有效的，在大电流负载下是没有任何差别的。因此，第二个名为“细节模型”的模型是通过假设不同组件间不同温度分布开发出来的，这些组件分别为电极、互连器、燃料和空气通道。需要注意的是，用于这项工作的长期细节模型被创造了，由于其相对于集总模型来说具有更详细的结构。然而，简化地假设，本章介绍的两种模型没有第 7 章介绍的管式 SOFC 详细。本章介绍的模型将要用来和 Padullés 等人研究的单一物质模型比较（2000）。

10.2　几何学

不同种类的平板型 SOFC 是基于形状、流量设定或歧管结构进行设计的。在大多数情况下，SOFC 被构建成矩形形状，气流方向有并流、逆流或横向流。电池阵列可以被串联、并联或基于要求的两种组合下连接。除了结构，燃料电池也可以根据阳极和阴极的气流来分类，即并流或逆流。在结构支撑的基础上，燃料电池主要可分为阳极、阴极或电解质支撑组装。歧管，在另一方面，可以是内部或外部的。一个典型的 SOFC 的安装范围可以从几千瓦的便携式电子产品到几兆瓦的大型公用事业规模的电站。Fuel Cell Hand book（燃料电池手册）（2004）提供了一个 SOFC 系统的世界范围制造商列表。

如前面所述，板式燃料电池的主要组成部分包括阳极、阴极和电解质。电解质被夹在阳极和阴极之间，作为反应物之间的直接接触，防止直接气化反应的一个障碍。反应的燃料和空气在电极的两侧流动。几个这样的单个电池被堆叠在一起，通过互连器串联连接完成电路。更简单的设计和短路电流路径产生的结果将导致减少对电流流动的阻碍。这样的一个安排如图 10.1 所示。

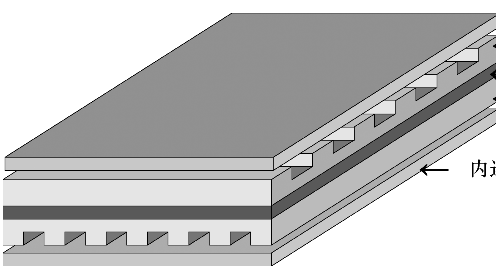


图 10.1　横向流式布置的板式 SOFC 堆栈的一个燃料电池单元视图

10.3　堆栈电压

在一个氢操作的燃料电池中，氢在阳极表面释放电子，这些电子穿过外电路，并与氧结合而产生氧离子。氧离子通过电解质到达阳极表面，在那里与氢离子结

合，形成水。该反应可概括如下：

$$\text{阳极：} H_2 + O^{2-} \rightarrow H_2O + 2e^- \tag{10.1}$$

$$\text{阴极：} \frac{1}{2}O_2 + 2e^- \rightarrow O^{2-} \tag{10.2}$$

电极的电位差 E 使得反应发生，通过电路的离子流 Nernst 方程可以表示为

$$E = E^{o} + \frac{RT_e}{2F}\ln\frac{p_{H_2}p_{O_2}^{0.5}}{p_{H_2O}} \tag{10.3}$$

式中，E^{o} 是标准电池电动势；p_{H_2}、p_{O_2}和 p_{H_2O}分别为氢、氧和蒸汽的分压；T_e 是电极在发生反应处的温度；R 和 F 分别是通用气体常数和法拉第常数。

继上述的法则，并考虑电阻、活化和浓度损失，N_0 串联堆叠的燃料电池堆栈电压可表示为

$$V_s = V_0 - \eta_{ohm} - \eta_{act} - \eta_{con} \tag{10.4}$$

开路电压 V_0 为

$$V_0 = N_0 E = N_0\left[E_0 + \frac{RT_e}{2F}\ln\frac{p_{H_2}p_{O_2}^{0.5}}{p_{H_2O}}\right] \tag{10.5}$$

在这里，标准电池电动势 E^{o} 展示了电池温度的线性关系，可以从 Fuel Cell Hand book（2004）的表2-3 中的实验数据中近似得出

$$E^{o} = 1.2586 - 0.000252T_e \tag{10.6}$$

由于通过电池材料的不同组成部分的离子流的电阻产生欧姆极化，损失可以表示为

$$\eta_{ohm} = r(T_e)I \tag{10.7}$$

电池电阻，$r(T_e)$ 是电池电极温度的函数，可以用 Steinhart- Hart 二阶方程表示：

$$r(T_e) = r_0\exp\left[\alpha\left(\frac{1}{T_e} - \frac{1}{T_0}\right)\right] \tag{10.8}$$

式中，r_0 是在 T_0 温度（K）时的内阻（Ω）；α 是常数。

由下式可以计算出活化和浓差极化（Braun，2002；Larminie 和 Dicks，2003；Sedghisigarchi 和 Feliachi，2004）：

$$\eta_{act} = a + b\log I \tag{10.9}$$

$$\eta_{con} = \frac{RT_e}{2F}\ln\left(1 - \frac{I}{I_L}\right) \tag{10.10}$$

式中，a 和 b 是 Tafel 常数及 Tafel 斜率；相应的，I_L 是极限电流。

使用上述模型，如果部分压力和电极温度是已知的，那么给定一个电流负载，就可以计算燃料电池的输出电压。指定 p_{H_2}、p_{O_2}、p_{H_2O}和 T_e 为 x_1、x_2、x_3 和 x_4，堆栈电压 V_s 为 y，该模型可以由下面的非线性代数方程表示：

$$y = V_s = N_0\left[1.2586 - 0.000252x_4 + \frac{Rx_4}{2F}\ln\frac{x_1x_2^{0.5}}{x_3}\right]$$

$$-r_0\exp\left[\alpha\left(\frac{1}{x_4}-\frac{1}{T_0}\right)\right]u_d-[a+b\log u_d]-\frac{Rx_4}{2F}\ln\left[1-\frac{u_d}{I_L}\right] \quad (10.11)$$

因此，堆栈电压的模型是 $y=f(x,u_d)$，电流负载 I 是干扰变量由 u_d 表示，其他参数是恒定的。以下各节提供了一个通过电池建模预测操作工艺条件 x_i 的详细内容。

10.4　质量守恒

下面是 Padulles 等人（2000）的方法，质量平衡方程在本部分描述。

燃料电池堆栈周围氢的组分平衡可以表示为

$$\frac{\mathrm{d}p_{H_2}}{\mathrm{d}t}=\frac{RT_e}{V_{an}}(\dot{n}_{H_2}^{in}-\dot{n}_{H_2}^{out}-\dot{n}_{H_2}^{r}) \quad (10.12)$$

式中，V_{an}是阳极室的体积（m^3）；$\dot{n}_{H_2}^{in}$、$\dot{n}_{H_2}^{out}$和$\dot{n}_{H_2}^{r}$分别是入口、出口和 H_2 的反应摩尔流率（$mol\ s^{-1}$）。

通过假定气体流路内的压力恒定，内部和外部的压力比足够大，以至于孔淤塞了，反应和出口氢气的流量可以表示为

$$\dot{n}_{H_2}^{r}=2K_rI \quad (10.13)$$

$$\dot{n}_{H_2}^{out}=K_{H_2}p_{H_2} \quad (10.14)$$

式中，$K_r=\frac{N_0}{4F}$；I 是堆叠电流（A）；N_0 是串联在堆栈里面的电池数；K_{H_2}是氢原子的阀摩尔常数。

定义 $\tau_{H_2}=\frac{V_{an}}{K_{H_2}RT_e}$，式（10.12）可以写成

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{H_2}=\frac{1}{\tau_{H_2}K_{H_2}}(\dot{n}_{H_2}^{in}-K_{H_2}p_{H_2}-2K_rI) \quad (10.15)$$

式中，K_{H_2}是常数；τ_{H_2}仅是电池温度的函数，可以表示成 $\tau_{H_2}=\frac{\tau_{H_2}^0T^0}{T_e}$，其中 $\tau_{H_2}^0=\tau_{H_2}\big|_{T_e=T^0}$。

因此，式（10.15）可以写为

$$\frac{\mathrm{d}p_{H_2}}{\mathrm{d}t}=\frac{T_e}{K_{H_2}\tau_{H_2}^0T^0}(\dot{n}_{H_2}^{in}-K_{H_2}p_{H_2}-2K_rI) \quad (10.16)$$

类似的，燃料电池组分 O_2 和 H_2O 的质量平衡，导致了下面的一组方程：

$$\frac{\mathrm{d}p_{O_2}}{\mathrm{d}t}=\frac{T_e}{K_{O_2}\tau_{O_2}^0T^0}(\dot{n}_{O_2}^{in}-K_{O_2}p_{O_2}-K_rI) \quad (10.17)$$

$$\frac{\mathrm{d}p_{H_2O}}{\mathrm{d}t}=\frac{T_e}{K_{H_2O}\tau_{H_2O}^0T^0}(\dot{n}_{H_2O}^{in}-K_{H_2O}p_{H_2O}+2K_rI) \quad (10.18)$$

同样，根据相同的状态空间模型表示法，上述方程可以概括为

$$\dot{x}_1 = \frac{x_4}{K_{H_2}\tau^0_{H_2}T^0}(u_1 - K_{H_2}x_1 - 2K_r u_d) \tag{10.19}$$

$$\dot{x}_2 = \frac{x_4}{K_{O_2}\tau^0_{O_2}T^0}(u_2 - K_{O_2}x_2 - K_r u_d) \tag{10.20}$$

$$\dot{x}_3 = \frac{x_4}{K_{H_2O}\tau^0_{H_2O}T^0}(u_3 - K_{H_2O}x_3 + 2K_r u_d) \tag{10.21}$$

其中反应物的流速$\dot{n}^{in}_{H_2}$、$\dot{n}^{in}_{O_2}$和$\dot{n}^{in}_{H_2O}$，在设计一个控制器时，可以根据可用性，指定为输入变量。

10.5 能量守恒

如前面提到的，集中和详细模型的主要区别是温度动态模型已经开发的方式。燃料电池不同组件的温度分布在详细模型中被假定为不同的，不同组件表现出的能量平衡在图 10.2 中表示出来。然而，在集中模型中，假定整个燃料电池堆栈的温度是均匀分布的，因此，围绕整个电池堆栈，可以计算一个整体能量平衡表。

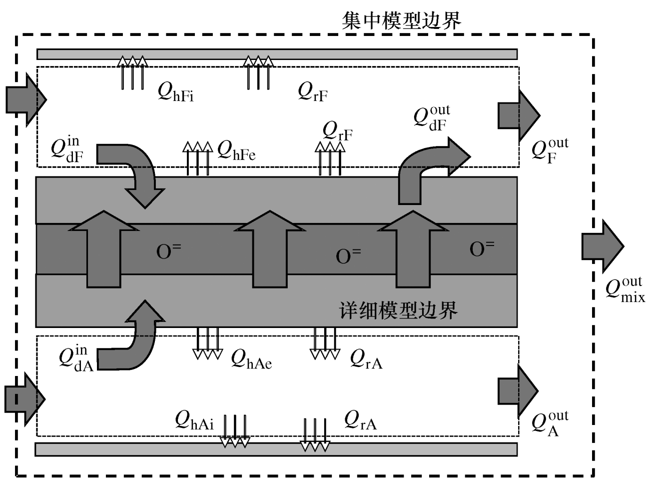


图 10.2 集中和详细模型的能量平衡边界层

除了温度分布以外，以下常见的假设是针对集中和详细模型的：

1）理想气体定律适用于两种反应物和排放物；

2）反应物和流出物的气体的热容量可以忽略不计；

3）散失到周围环境的热量可以忽略不计。

10.5.1 集中模型

在集中堆叠模型中，假定不存在燃料电池内部的温度变化，这意味着燃料电池

用电极、互连器和通道内气体的所有组件，在任何情况下，具有相同的温度。电池温度的动态模型和电极温度是相同的 T_e，可以通过假设围绕整个燃料电池堆栈进行能量平衡的计算得到：

$$m_e \overline{C}_p \frac{dT_e}{dt} = \sum \dot{n}_i^{in} \int_{T_{ref}}^{T_{in}} C_{p,i}(T)\,dT - \sum \dot{n}_i^{out} \int_{T_{ref}}^{T_e} C_{p,i}(T)\,dT - \dot{n}_{H_2}^{r} \Delta \hat{H}_r^o - V_s I \tag{10.22}$$

式中，m_e 和 $\overline{C}_p$ 是燃料电池材料除气体外，即电极和互连器的质量和平均比热；$C_{p,i}$是第 i 个燃料或系统中气体的比热；$\Delta \hat{H}_r^o$ 是反应的比热容；V_s 是堆叠电压。

考虑工作温度范围内的气体比热容是恒定的，上述等式可以简化为

$$m_e \overline{C}_p \frac{dT_e}{dt} = \sum \dot{n}_i^{in} C_{p,i}^{in}(T_{in} - T_{ref}) - \sum \dot{n}_i^{out} C_{p,i}^{out}(T_e - T_{ref}) - 2K_r I \Delta \hat{H}_r^o - V_s I \tag{10.23}$$

考虑到状态变量，式（10.23）可以写成

$$\dot{x}_4 = \frac{1}{m_e \overline{C}_p}\left[\sum \dot{n}_i^{in} C_{p,i}^{in}(T_{in} - T_{ref}) - \sum \dot{n}_i^{out} C_{p,i}^{out}(x_4 - T_{ref}) - 2K_r \Delta \hat{H}_r^o u_d - f(x_i, u_d) u_d\right] \tag{10.24}$$

根据不同燃料电池系统的设计，作为一种可能的操作变量和控制器配置的可用性变量，式（10.24）中反应物的入口温度 T_{in}是可操作的或可干扰变量。

10.5.2　详细模型

因为它会在后面部分显示，在更高的电流负载的集中模型中，燃料电池中温度分布均匀的假设是不合理的。因此，本节将介绍一个更详细的模型，假设不同的电池组件有不同的温度分布，因此，相对于“集中”模型来说是“详细的”。主要假设使它区别于集中模型的是：

1）电极、互连器、燃料和空气侧的气体之间的温度不同；

2）燃料电池内部通道的气体处于理想混合的状态，使得燃料和空气的出口温度与燃料和空气管道的温度相同；

3）在轴向方向没有温度变化。

基于这些假设，电极、互连器、燃料气体通道里质量和流量平衡，燃料电池中的空气通道在图 10.2 中描绘。在这个模型中，只有温度分布被假定为是不同的。电池周围的各种平衡和在 10.4 节描述的一样，电池电压的计算也和 10.3 节一样。

10.5.2.1　电极周围的能量平衡

电极控制体积由阳极、电解质和阴极组成。尽管电化学反应发生在电解质表面附近的阳极和阴极，因此，在表面区域的垂直方向有温度的变化，由于电极厚度非常小，温度可以假定是恒定的。沿流动方向的温度变化也假定为恒定的，因此，电极温度 T_e 的动态变化，可以表示成扩散、对流、辐射和反应传热的函数：

$$\rho_e A_c \Delta w_e \overline{C}_{pe} \frac{\mathrm{d}T_\mathrm{e}}{\mathrm{d}t} = (Q_{\mathrm{dF}}^{\mathrm{in}} - Q_{\mathrm{dF}}^{\mathrm{out}}) + (Q_{\mathrm{dA}}^{\mathrm{in}} - Q_{\mathrm{dA}}^{\mathrm{out}})$$

$$-(Q_{h\mathrm{Fe}} + Q_{\mathrm{rF}} + Q_{h\mathrm{Ae}} + Q_{\mathrm{rA}}) - Q^r - \dot{W} \quad (10.25)$$

式中，A_c、Δw_e、ρ_e 和 $\overline{C}_{pe}$ 分别为电化学表面积、厚度、密度和电极材料的平均比热。

式（10.25）燃料侧扩散热可以表示为

$$Q_{\mathrm{dF}}^{\mathrm{in}} = \dot{n}_{\mathrm{H_2}}^{\mathrm{r}} \int_{T_{\mathrm{ref}}}^{T_\mathrm{F}} C_{p,\mathrm{H_2}}(T)\,\mathrm{d}T = 2K_\mathrm{r} I C_{p,\mathrm{H_2}}(T_\mathrm{F} - T_{\mathrm{ref}}) \quad (10.26)$$

$$Q_{\mathrm{dF}}^{\mathrm{out}} = \dot{n}_{\mathrm{H_2O}}^{\mathrm{r}} \int_{T_{\mathrm{ref}}}^{T_\mathrm{F}} C_{p,\mathrm{H_2O}}(T)\,\mathrm{d}T = 2K_\mathrm{r} I C_{p,\mathrm{H_2O}}(T_\mathrm{F} - T_{\mathrm{ref}}) \quad (10.27)$$

式中，$\dot{n}_{\mathrm{H_2}}^{\mathrm{r}} = \dot{n}_{\mathrm{H_2O}}^{\mathrm{r}} = 2K_\mathrm{r}I$；$T_\mathrm{F}$ 是耗尽燃料的出口温度。

式（10.25）中，电极内、外部的空气侧传热可以表述为

$$Q_{\mathrm{dA}}^{\mathrm{in}} = \dot{n}_{\mathrm{O_2}}^{\mathrm{r}} \int_{T_{\mathrm{ref}}}^{T_\mathrm{A}} C_{p,\mathrm{O_2}}(T)\,\mathrm{d}T = K_\mathrm{r} I C_{p,\mathrm{O_2}}(T_\mathrm{A} - T_{\mathrm{ref}}) \quad (10.28)$$

$$Q_{\mathrm{dA}}^{\mathrm{out}} = 0 \quad (10.29)$$

式中，$\dot{n}_{\mathrm{O_2}}^{\mathrm{r}} = K_\mathrm{r}I$；$T_\mathrm{A}$ 是耗尽空气的出口温度。

式（10.25）中，电极和燃料/空气气体之间的对流热转移可以写成

$$Q_{h\mathrm{Fe}} = h_{\mathrm{Fe}} A_c (T_\mathrm{e} - T_F) \quad (10.30)$$

$$Q_{h\mathrm{Ae}} = h_{\mathrm{Ae}} A_c (T_\mathrm{e} - T_\mathrm{A}) \quad (10.31)$$

燃料侧和空气侧的传热系数 h_{Fe} 和 h_{Ae} 也是 T_F 和 T_A 的函数，通过每个实例经验性评估。

燃料侧电极和互连器之间的辐射传热方面，式（10.25）中空气侧可以表示为

$$Q_{\mathrm{rF}} = \frac{\sigma A_\mathrm{c}(T_\mathrm{e}^4 - T_\mathrm{i}^4)}{1/\epsilon_\mathrm{a} + 1/\epsilon_\mathrm{i} - 1} \quad (10.32)$$

$$Q_{\mathrm{rA}} = \frac{\sigma A_\mathrm{c}(T_\mathrm{e}^4 - T_\mathrm{i}^4)}{1/\epsilon_\mathrm{c} + 1/\epsilon_\mathrm{i} - 1} \quad (10.33)$$

式中，$\sigma = 5.6704 \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$，是玻尔兹曼常数；$T_\mathrm{i}$ 是互连器的温度（K）；ϵ_a、ϵ_c 和 ϵ_i 是阳极、阴极和互连器的材料的发射率常数。

在这种情况下，辐射的几何因子假定为 1。

最后，产生的热量和燃料电池所产生的电力可以表示为

$$Q^\mathrm{r} = -\dot{n}_{\mathrm{H_2}}^{\mathrm{r}} \Delta \hat{H}_\mathrm{r}^\mathrm{o} = -2K_\mathrm{r} I \Delta \hat{H}_\mathrm{r}^\mathrm{o} \quad (10.34)$$

$$\dot{W} = V_\mathrm{s} I$$

考虑到状态空间变量，模型可以表示为

$$\dot{x}_4 = \frac{1}{\rho_e A_c \Delta w_e \overline{C}_{pe}} [2K_\mathrm{r}(C_{p,\mathrm{H_2}} - C_{p,\mathrm{H_2O}})(x_6 - T_{\mathrm{ref}})u_\mathrm{d} + K_\mathrm{r} C_{p,\mathrm{O_2}}$$

$$(x_7 - T_{\mathrm{ref}})u_\mathrm{d} - A_\mathrm{c}(h_{\mathrm{Fe}}(x_4 - x_6) + h_{\mathrm{Ae}}(x_4 - x_7))$$

$$-\sigma A_{c}\left(\frac{1}{1/\epsilon_{a}+1/\epsilon_{i}-1}+\frac{1}{1/\epsilon_{c}+1/\epsilon_{i}-1}\right)$$

$$(x_4^4-x_5^4)+2K_{r}\Delta\hat{H}_{r}^{o}u_{d}-f(x_{i},u_{d})u_{d}\Big] \tag{10.35}$$

这里电极温度 T_e、互连器温度 T_i、燃料出口温度 T_F、空气出口温度 T_A 分别记做 x_4、x_5、x_6 和 x_7。

10.5.2.2　互连器的能量平衡

互连器的动态温度 T_i 可以表示为以下等式：

$$\rho_{i}A_{i}\Delta w_{i}\,\overline{C}_{pi}\frac{dT_{i}}{dt}=Q_{hFi}+Q_{rF}+Q_{hAi}+Q_{rA} \tag{10.36}$$

式中，ρ_i、A_i、Δw_i、$\overline{C}_{pi}$和 T_i 分别为互连器材料的密度、面积、厚度、比热容和温度。

互连器和燃料/空气之间的对流热转移表示如下：

$$Q_{hFi}=h_{Fi}A_{i}(T_{F}-T_{i}) \tag{10.37}$$

$$Q_{hAi}=h_{Ai}A_{i}(T_{A}-T_{i}) \tag{10.38}$$

式中，h_{Fi}和 h_{Ai}分别为毗邻互连器表面的燃料侧和空气侧的对流传热系数，这是在每个实例经验中估计出来的。

式（10.36）中辐射传热方面表示在式（10.32）和式（10.33）中。对于板式 SOFC，区域 A_i 可以近似为电化学表面积 A_c。

考虑到状态变量，模型可以表示为

$$\dot{x}_{5}=\frac{1}{\rho_{i}A_{i}\Delta w_{i}\,\overline{C}_{pi}}\Big[h_{Fi}A_{i}(x_{6}-x_{5})+h_{Ai}A_{i}(x_{7}-x_{5})+\sigma A_{c}$$

$$\left(\frac{1}{1/\epsilon_{a}+1/\epsilon_{i}-1}+\frac{1}{1/\epsilon_{c}+1/\epsilon_{i}-1}\right)(x_4^4-x_5^4)\Big] \tag{10.39}$$

10.5.2.3　燃料侧的能量平衡

出口燃油温度的动态模型，T_F 可以表示为

$$\sum\frac{p_{i}V_{channel}C_{pi}^{out}}{RT_{F}}\frac{dT_{F}}{dt}=(Q_{F}^{in}-Q_{F}^{out})+(Q_{dF}^{out}-Q_{dF}^{in})+(Q_{hFe}-Q_{hFi}) \tag{10.40}$$

燃料电池的吸热 Q_F^{in} 和放热 Q_F^{out} 可以表示为

$$Q_{F}^{in}=\sum\dot{n}_{i}^{in}\int_{T_{ref}}^{T_{F}^{in}}C_{P,i}(T)\,dT=\sum\dot{n}_{i}^{in}C_{P,i}(T_{F}^{in}-T_{ref}) \tag{10.41}$$

$$Q_{F}^{out}=\sum\dot{n}_{i}^{out}\int_{T_{ref}}^{T_{F}}C_{P,i}(T)\,dT=\sum\dot{n}_{i}^{out}C_{P,i}(T_{F}-T_{ref}) \tag{10.42}$$

式中，T_F^{in} 是进入 SOFC 的燃料的温度。

式（10.40）中扩散和对流的传热 Q_{dF}^{in}、Q_{dF}^{out}、Q_{hFe}和 Q_{hFi}分别在式（10.26）、式（10.27）、式（10.30）和式（10.37）中表示。

考虑到状态变量，这个模型可以概括为

$$\dot{x}_6=\frac{1}{\sum\frac{p_i V_{\mathrm{channel}}C_{pi}^{\mathrm{out}}}{Rx_6}}\Big[\sum \dot{n}_i^{\mathrm{in}}C_{P,i}(T_F^{\mathrm{in}}-T_{\mathrm{ref}})u_i-\sum \dot{n}_i^{\mathrm{out}}C_{P,i}(x_6-T_{\mathrm{ref}})-2K_r(C_{p,H_2}-C_{p,H_2O})(x_6-T_{\mathrm{ref}})u_d+h_{Fe}A_c(x_4-x_6)-h_{Fi}A_i(x_6-x_5)\Big] \tag{10.43}$$

类似于集总模型，根据在燃料电池系统的配置和控制器的设计，燃料入口温度 T_F 可能被设计为可调节或干扰的变量。

10.5.2.4　空气侧周围的能量平衡

空气出口温度 T_A 的动态模型可以表示为

$$\sum\frac{p_i V_{\mathrm{channel}}C_{pi}^{\mathrm{out}}}{RT_A}\frac{\mathrm{d}T_A}{\mathrm{d}t}=(Q_A^{\mathrm{in}}-Q_A^{\mathrm{out}})+(Q_{dA}^{\mathrm{out}}-Q_{dA}^{\mathrm{in}})+(Q_{hAe}-Q_{hAi}) \tag{10.44}$$

式中，Q_A^{in} 和 Q_A^{out}分别为通过空气流的燃料电池的流入和流出的热量，它们可以表示为

$$Q_A^{\mathrm{in}}=\sum \dot{n}_i^{\mathrm{in}}\int_{T_{\mathrm{ref}}}^{T_A^{\mathrm{in}}}C_{P,i}(T)\,\mathrm{d}T=\sum \dot{n}_i^{\mathrm{in}}C_{P,i}(T_A^{\mathrm{in}}-T_{\mathrm{ref}}) \tag{10.45}$$

$$Q_A^{\mathrm{out}}=\sum \dot{n}_i^{\mathrm{out}}\int_{T_{\mathrm{ref}}}^{T_A}C_{P,i}(T)\,\mathrm{d}T=\sum \dot{n}_i^{\mathrm{out}}C_{P,i}(T_A-T_{\mathrm{ref}}) \tag{10.46}$$

式（10.44）中的扩散和对流的传热 Q_{dA}^{in}、Q_{dA}^{out}、Q_{hAe}和 Q_{hAi}分别在式（10.28）、式（10.29）、式（10.31）和式（10.38）中表示。

考虑到状态变量，模型可以写成

$$\dot{x}_7=\frac{1}{\sum\frac{p_i V_{\mathrm{channel}}C_{pi}^{\mathrm{out}}}{Rx_7}}\Big[\sum \dot{n}_i^{\mathrm{in}}C_{P,i}(T_A^{\mathrm{in}}-T_{\mathrm{ref}})u_i-\sum \dot{n}_i^{\mathrm{out}}C_{P,i}(x_7-T_{\mathrm{ref}})-K_rC_{p,O_2}(x_7-T_{\mathrm{ref}})u_d+h_{Ae}A_c(x_4-x_7)-h_{Ai}A_i(x_7-x_5)\Big] \tag{10.47}$$

类似于燃料入口温度，空气入口温度 T_A^{in} 是一个可以调节的或是干扰的变量。

10.6　仿真

在上一节衍生出来的模型已经被列作一组非线性常微分方程，这种模式适合于仿真燃料电池在不同的操作条件下的行为。然而，这个模型参数需要从实验和/或文献数据仿真和控制器的设计中得出。这些数据大部分都摘录自 Braun（2002）、Padullés 等人（2000）、Perry（1997）、Sedghisigarchi 和 Feliachi（2004）以及 Fuel Cell Hand book（2004）。有些模型参数不可用，但已在合理的范围内得出。表 10.1 包含了由串联的 384 个电池堆栈而成的板式 SOFC 的参数。这些模型现在可以用来仿真不同操作条件下的稳态和动态行为。

表 10.1　模型参数

参　数	值	单　位
电池数量，N_0	384	
电池面积，A_c	100	cm^2
K_{H_2}	8.43×10^{-4}	kmol atm. s^{-1}
K_{H_2O}	2.81×10^{-4}	kmol atm. s^{-1}
K_{O_2}	2.52×10^{-3}	kmol atm. s^{-1}
$\tau_{H_2}^*$	26.1	s
$\tau_{H_2O}^*$	78.3	s
$\tau_{O_2}^*$	2.91	s
T^*	1000	℃
电阻常量，r_0	0.126	Ω
电阻梯度，α	−2870	
电极厚度，h_e	0.25	mm
互连器厚度，h_i	1.5	mm
电极密度，ρ_e	6.6	g cm^{-3}
互连器密度，ρ_i	6.11	g cm^{-3}
比热，$\bar{C}_{pe}$、$\bar{C}_{pi}$、$\bar{C}_{ps}$	0.4	J g. K^{-1}
反应热，$\Delta\hat{H}_r^\circ$	-0.2418×10^9	J $kmol^{-1}$
辐射率，ϵ_a、ϵ_c、ϵ_i	0.9	

10.6.1　稳态响应

和文献相比，燃料电池的稳态响应被用来模拟细节和集中模型。该模型被用于产生功率、电压以及当保持燃料、空气的速度和温度保持恒定条件下的响应电流负载的电池温度。在这个模拟中，只有欧姆损耗被视为与单一物质模型的结果相比较（Padullés 等人，2000）。

图 10.3a、b 显示了详细、集中和单一物质模型下的板式 SOFC 的稳态功率和电压输出，单一物质模型引自 Padullés 等人（2000）的模型。从图中可以看出，显而易见的是，从提出的模型中得出的功率和电压是相符的。然而单一物质动态模型预测较高的电池终端的功率和电压。事实上，根据基于电池工作温度的单一物质模型预测的堆栈电压和功率不是恒定的，它随操作条件不同而变化。

从图 10.3 中还可以看出，即使所有的模型预测中的分压几乎是相同的，但不同模型之间的电池温度有很大的差异。因此，导致了不同的电压和功率，堆栈电压是电池温度的函数。在更大的电流负载条件下，温度的影响更为突出。因此，重要的是根据目标，选择一个合适的模型。例如，在计算能力有限的条件下，在设计控

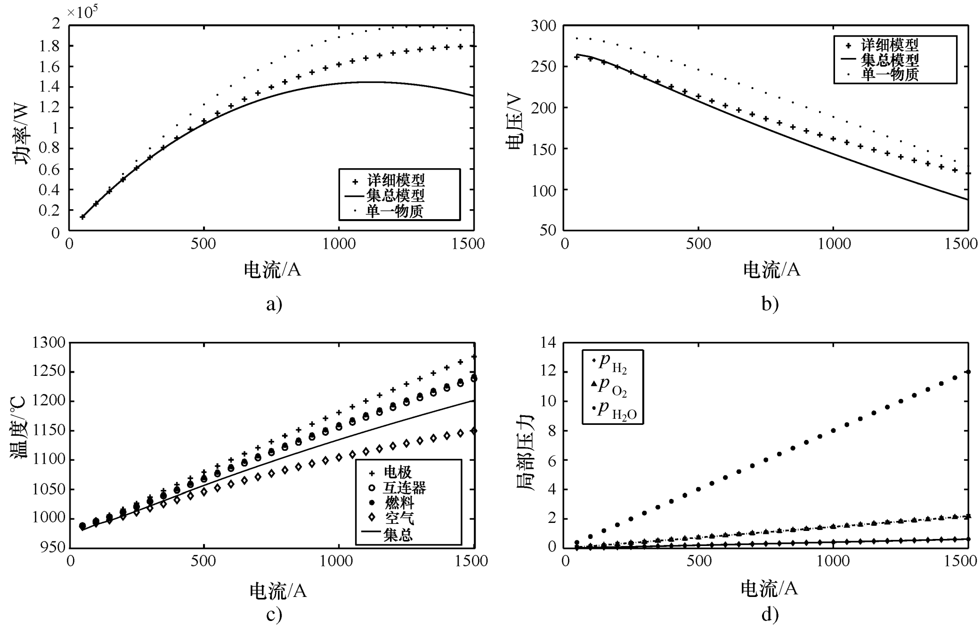


图 10.3　详细、集中和单一物质模型下的板式 SOFC 的稳态曲线

a）功率-电流　b）电压-电流　c）温度-电流　d）局部压力-电流

制器时，集中模型可以运用，但可能不适合用于评价其性能，特别是在一个燃料电池系统。

10.6.2　动态响应

通过改变负载、燃料和空气的速率和温度，也对这两种模型进行了动态模拟。在保持燃料和空气速度不变的条件下，当前需求电流变化的影响如图 10.4 所示。在这种情况下，反应物的入口温度也被保持恒定。因为模型是在开环状态下进行的仿真，也就是说，没有任何控制行为，燃料利用率一直保持在一个相对较低的值，以保障燃料电池对抗燃料不足。在所有条件下，与需求电流变化相关联，堆栈电压也有一个突降。这个突然的电压降来自式（10.4）中堆栈电压的欧姆损耗 $r(T)I$。这种逆反应是一种独特的燃料电池的特征，需要特别注意的是，燃料电池操作期间需要保持电压的稳定。

图 10.5 和图 10.6 说明，当保持其他输入变量恒定的条件下，燃料和空气的速度和入口温度变化时，燃料电池的瞬间变化。在这些特殊情况下，燃料利用率一直保持 0.85 左右。由于电极温度是负载电流的函数，当负载增加时，电极温度也增加。电池温度的增加和内部电阻的欧姆损耗以及式（10.8）、式（10.6）中的标准电极电位相关。因此，根据欧姆损耗的幅度和标准电极电位，堆栈电压有收益或损失。

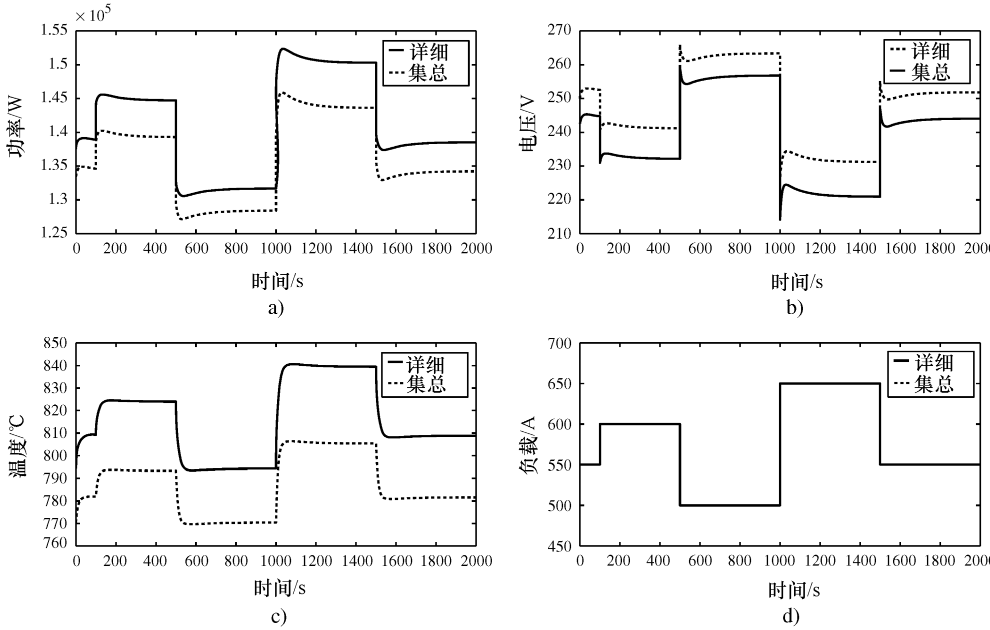


图 10.4　负载变化对功率、电压和电池温度变化的影响

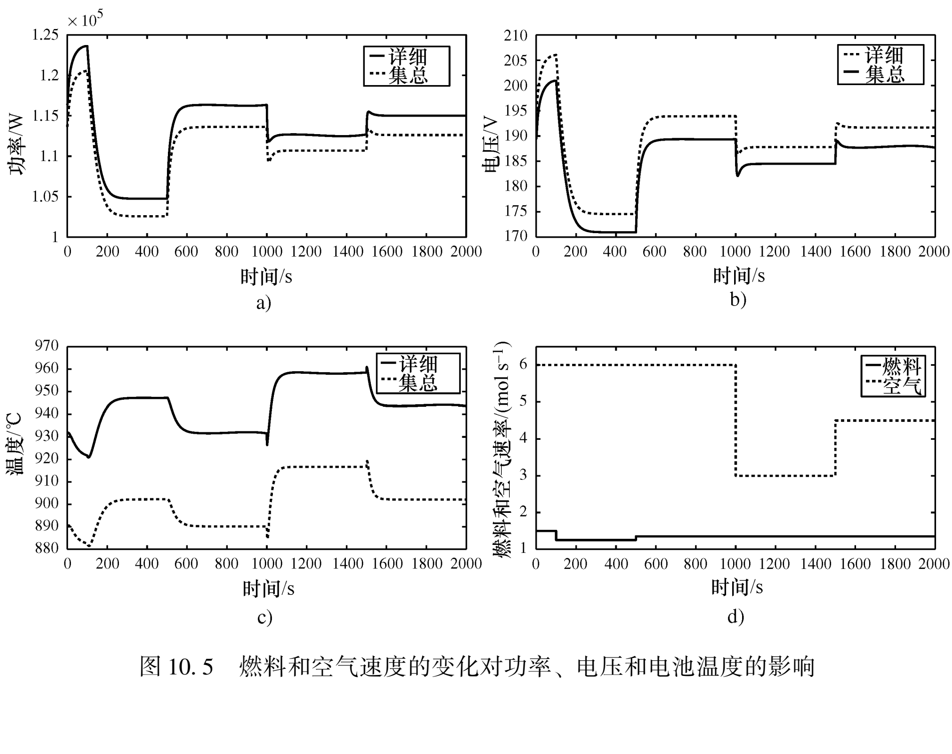


图 10.5　燃料和空气速度的变化对功率、电压和电池温度的影响

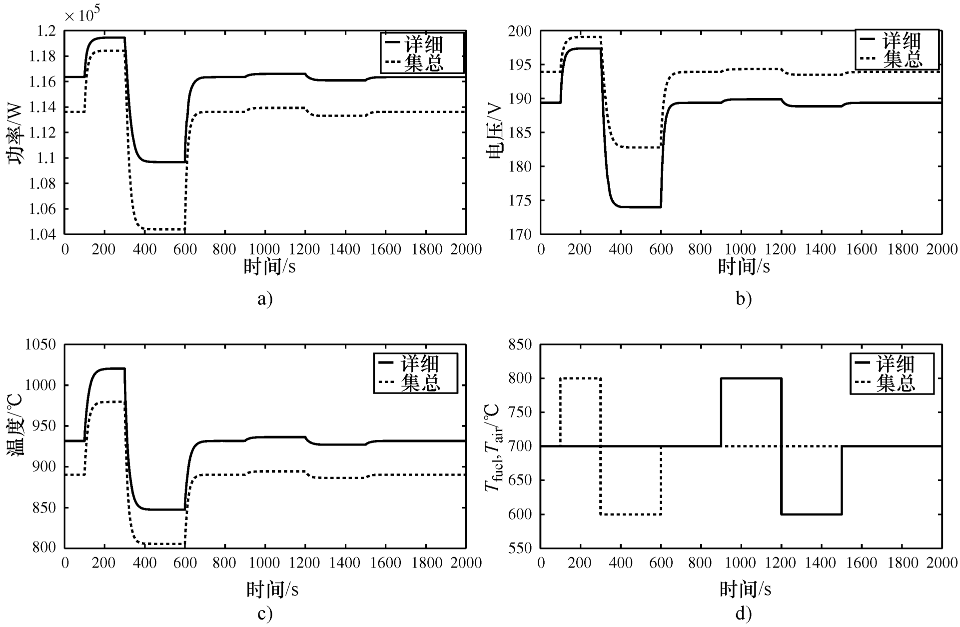


图 10.6　燃料和空气温度的变化对功率、电压和电池温度的影响

10.7　注解和参考文献

在本章，运用两个模型，集中和详细模型来代表 SOFC 的动态模型。零维模型可以很容易地用来预测稳态和动态响应，以及基于模型的控制器设计，如线性和非线性 MPC。此外，由于不是所有的状态都是可以衡量的，使用状态估计器，如扩展卡尔曼滤波（EKF）或无迹卡尔曼滤波（UKF），相同的模型可以用于估算状态。根据第 2 章中介绍的原则，使用燃料电池周围的质量和能量平衡，这些模型已制定了。化学过程，特别是板式 SOFC 的动态模型的详细理论可以在 Felder 和 Rousseau（1986）、Lienhard Ⅳ 和 Ⅴ（2005）和 Perry 等人（1997）以及 Murshed 等人（2007）的文献中找到。

第 11 章　板式 SOFC 系统动态建模

11.1　简介

在前面的章节中，独立的燃料电池已经被描述和建模。最终的模型由一组非线性微分方程组成，它可以很容易地用于仿真开环瞬态特性或设计基于模型的控制。SOFC 通常是被馈以氢气为燃料和空气作为氧的来源。这种单机燃料电池可以用于驱动小型设备，但可能不适合于大的功率生产设备。在后一种情况下，燃料电池与预处理和后处理单元相关联，以使其供给天然气或其他富含氢气的燃料而不是纯氢气。同样重要的是，从流出物中回收废气以增加燃料电池的整体效率。一个这样的 SOFC 系统如图 11.1 所示，其中包括一个重整器，热交换器用于在燃料和空气加入之前对它们预加热，再分别送入重整器和燃料电池堆栈，燃烧器用于燃烧排放物。通常，燃料电池系统包括利用废气产生辅助动力的涡轮机、燃料处理装置以及功率调节装置。这些组件是燃料电池系统的完整组成部分，并被称作 BOP。

在本章中，上述各部件的集总模型都会呈现出，然后与燃料电池模型集成来模拟燃料电池系统的动态行为。类似于独立的燃料电池，该燃料电池系统的组件模型是由一组可用于模拟系统的常微分方程描述的。该模型也可以用来设计整个系统的控制器，在后面的内容中会讨论。

11.2　燃料电池系统

在如图 11.1 所示的例子中，用甲烷作为燃料电池系统的燃料。天然气或富烃气体也可以被用作燃料电池系统的燃料，这需要设计相关燃料预处理器。

为简单起见，甲烷被作为燃料，它可以通过重整和水煤气变换反应很容易地转化为氢。在进入重整装置前，通过从燃料电池流出物中的热量，首先将燃料预热至所需的温度。产物气体通过气体转换器进入燃料电池堆栈的阳极室。压缩空气也要通过另一个热交换器对其进行预加热，通过燃料热交换器的热出口，最后被送到燃料电池堆栈的阴极室。阳极室的氢和阴极室的氧在电极处参与电化学反应来产生能量，正如前面的章节所讨论的那样。

耗尽的燃料和燃料电池堆栈中的空气被送入一个燃烧器，从燃料电池的流出物中产生热量。燃烧器的排气的一部分被送到燃料和空气预热器，其余部分用来产生蒸汽，蒸汽可以用于转化反应。可供选择的，流出物可以直接用于在锅炉中

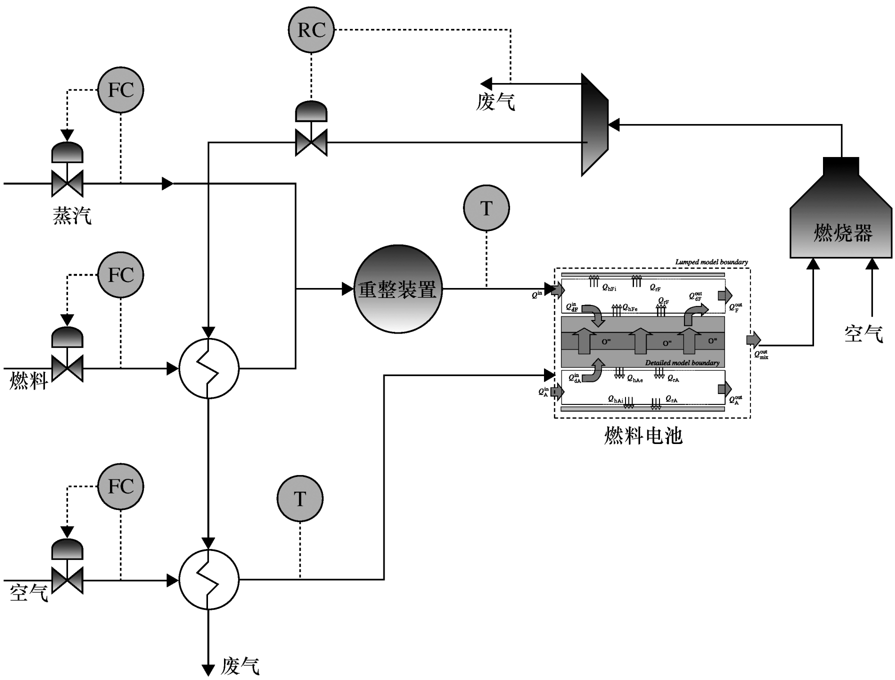


图 11.1 SOFC 系统的热交换器、重整装置、燃烧器和压缩机

产生蒸汽，这可以作为产生能量的蒸汽的来源通过蒸汽涡轮机。为保持简单模型，本章不包括蒸汽锅炉模型。然而，利用相同的概念，锅炉和涡轮机可以建模，模拟燃料电池的动态行为。以下各节会对燃料电池系统的各个组成部分的造型进行简要说明。

11.2.1 燃料和空气热交换

在这个例子中，两个热交换器被认为是用于对燃料电池系统的燃料和空气预加热。这两个热交换器被假定为逆流双管换热器。燃烧器产生的废气的一部分热流被供给到外管，冷流被送入内管逆流。假定热交换器以恒定的压力工作（一个温和的假设），则热交换器的总体热模型可表示如下：

$$\rho C_p A \frac{\partial T}{\partial t} = -\rho C_p v A \frac{\partial T}{\partial z} + \pi D Q \tag{11.1}$$

式中，ρ 是气体密度；C_p 是气体的定压比热容；A 是横截面面积；Q 是每单位面积的热传导率基于热转移面积 πD。

如式（11.1）所示，热交换器模型是一个偏微分方程，所以很难设计、控制或仿真与其他组分相结合的动态响应。因此，通过将沿流动方向上的热交换器送入 n 个节点，该模型被转换成一组常微分方程，如图 11.2 所示。假定每个节点作为一个连续搅拌釜式加热器，以便每个节点内出口的气体的温度是相同的。然后，第 i 个节点的能量平衡为管（冷气体）和壳（热气体）可以写为

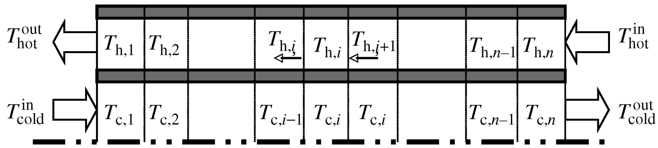


图 11.2　热交换器沿长度分成 n 个节点

管：

$$\rho_c C_{pc} A_{cx} \frac{dT_c(i)}{dt} = -\rho_c C_{pc} v_c A_{cx} \frac{T_c(i) - T_c(i-1)}{\Delta z} + \pi D_0 U_0 (T_h(i) - T_c(i)) \tag{11.2}$$

壳：

$$\rho_h C_{ph} A_{hx} \frac{dT_h(i)}{dt} = -\rho_h C_{ph} v_h A_{hx} \frac{T_h(i+1) - T_h(i)}{\Delta z} - \pi D_0 U_0 (T_h(i) - T_c(i)) \tag{11.3}$$

式中，下角 c 和 h 分别代表冷侧和热侧的能量平衡；A_{cx} 和 A_{hx} 分别代表冷侧和热侧的流动横截面积；Δz 是每个节点的长度；D_0 是管的外径；U_0 为整体散热传热系数，表示为

$$U_0 = \left(\frac{r_0}{r_i h_i} + \frac{r_0 \ln(r_0/r_i)}{k_w} + \frac{1}{h_i} \right)^{-1} \tag{11.4}$$

常微分方程的集中换热器模型的总数在此示例中为 $2n$。

11.2.2　重整装置

来自燃料热交换器预热的甲烷和单独流动的蒸汽被送到重整装置，其中接下来的吸热反应在甲烷和 H_2O 之间发生：

重整反应：

$$CH_4 + H_2O \rightarrow CO + 3H_2,\ \Delta \hat{H}_R^o = 206.1 kJ\ mol^{-1} \tag{11.5}$$

如果重整反应速率由 $\dot{r}_R$ 表示，那么大规模的反应物和结余产品如下：

$$\dot{n}_{CH_4}^{out} = \dot{n}_{CH_4}^{in} - \dot{r}_R \tag{11.6}$$

$$\dot{n}_{H_2O} = \dot{n}_{H_2O}^{in} - \dot{r}_R \tag{11.7}$$

$$\dot{n}_{CO} = \dot{r}_R \tag{11.8}$$

$$\dot{n}_{H_2} = 3\dot{r}_R \tag{11.9}$$

式中

$$\dot{r}_R = k_0 P_{CH_4} \exp\left(-\frac{E_A}{RT_R} \right) \cdot A_{RX} mol | s^{-1} \tag{11.10}$$

$$k_0 = 4274\text{mol m}^{-2}\text{ s bar} \tag{11.11}$$

$$E_A = 82\text{kJ mol}^{-1} \tag{11.12}$$

将 CO 和 H_2O 混合，然后取一部分参与可逆放热变换反应来产生更多的氢。

水煤气变换反应：

$$CO + H_2O \rightleftharpoons CO_2 + H_2 \quad \Delta\hat{H}_S^o = -41.15\text{kJ mol}^{-1} \tag{11.13}$$

如果上述反应的反应和平衡常数的范围分别由 ξ 和 K_S 表示，那么对反应物和产物组分的质量平衡如下：

$$\dot{n}_{H_2}^{out} = \dot{n}_{H_2} + \xi \tag{11.14}$$

$$\dot{n}_{CO_2}^{out} = \xi \tag{11.15}$$

$$\dot{n}_{CO}^{out} = \dot{n}_{CO} - \xi \tag{11.16}$$

$$\dot{n}_{H_2O}^{out} = \dot{n}_{H_2O} - \xi \tag{11.17}$$

式中

$$K_S(T_R) = \frac{\xi(\dot{n}_{H_2} + \xi)}{(\dot{n}_{CO} - \xi)(\dot{n}_{H_2O} - \xi)} \tag{11.18}$$

平衡常数 K_S 可以表示为温度的函数（Bove 等人，2005）：

$$K_S(T_R) = \exp\left(\frac{5693.5}{T_R} + 1.077\ln T_R + 5.44\times10^{-4}T_R - 1.125\times10^{-7}T_R^2 - \frac{49170}{T_R^2} - 13.148\right) \tag{11.19}$$

现在 K_S 是已知的，反应的程度可以通过求解式（11.18）计算。围绕上述重整的能量平衡可以被写为

$$\rho V_R C_p \frac{dT_R}{dt} = \dot{n}_{CH_4}^{in}\int_{T_{ref}}^{T_{CH_4}^{in}} C_{p,CH_4}(T)\,dT + \dot{n}_{H_2O}^{in}\int_{T_{ref}}^{T_{H_2O}^{in}} C_{p,H_2O}(T)\,dT - \sum \dot{n}_i^{out}\int_{T_{ref}}^{T_R} C_{p,i}(T)\,dT + r_R\Delta\hat{H}_R^o + \xi\Delta\hat{H}_S^o \tag{11.20}$$

式中，T_R 是重整装置的操作温度。因此，被送至燃料电池的流出物可以使用上面设置的常微分方程模拟重整。

11.2.3　燃烧器

最后，将燃料电池的未反应的燃料和空气一起燃烧用于热恢复。在这个阶段，为它提供额外的氧气是很重要的，以使燃料电池中未反应的燃料被消耗掉。这是通过用另一个气流进入燃烧器进行的。假设燃烧器工作在恒定的压力且有理想气体在腔室内部混合，以便使燃烧器的出口和内部有相同的温度，燃烧器周围的能量平衡可表示为

$$\rho V_B \bar{C}_p \frac{dT_B}{dt} = \sum \dot{n}_{iF}^{in} \int_{T_{ref}}^{T_F} C_{p,i}(T)dT + \sum \dot{n}_{iA}^{in} \int_{T_{ref}}^{T_A} C_{p,i}(T)dT - \sum \dot{n}_{iB}^{out} \int_{T_{ref}}^{T_B} C_{p,i}(T)dT - \dot{n}_{H_2}^{in} \Delta\hat{H}_{r,H_2}^{o} - \dot{n}_{CH_4}^{in} \Delta\hat{H}_{r,CH_4}^{o} - \dot{n}_{CO}^{in} \Delta\hat{H}_{r,CO}^{o} \quad (11.21)$$

式中，T_B 是燃烧器温度；下标 F、A 和 B 分别代表燃料、空气和燃烧器；比热 $\bar{C}_p$ 和密度 ρ 也被根据每一个计算实例通过经验计算出来。

11.3　连同电容器的 SOFC

如前面所述，一个独立的燃料电池的模型可以简要地表示为

$$\dot{x} = f(x, \dot{n}_i, I_d, V_s) \quad (11.22)$$

$$V_s = V_0(x) - r(x) I_d \quad (11.23)$$

式中，x 包含描述的分压和 SOFC 的状态模型的温度；n_i 代表燃料或空气的吸入流速；I_d 是穿过燃料电池需求的电流负载；V_s 是燃料电池产生的堆栈电压。

为简单起见，在上述输出方程中只考虑欧姆损耗。对于一个独立的燃料电池，燃料电池电流 I_{fc} 等于需求电流 I_d。从输出方程显而易见的是，需求电流的突然变化与瞬态响应所表现出的瞬时电压相关联（见图 10.4）。堆栈电压的瞬时变化是不可避免的，无论用多么先进的类型控制，这是因为在入口的燃料和空气流速等操纵变量的制约。为了避免突然地电压损失和电气设备的损坏，足够容量的电容器可以和燃料电池并联使用作为辅助电源，这被描绘在图 11.3 中。电容器的优点可以直

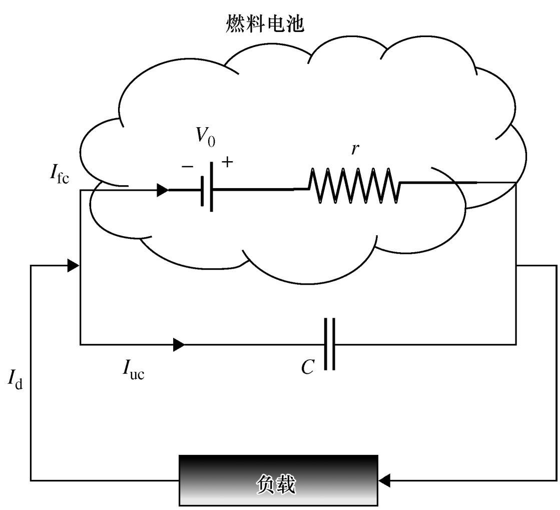


图 11.3　与电容器并联的 SOFC

观地看出：在有需求电流突然变化的时候，该电容将分担负载并提供额外的电源。因此，堆栈电压不会突然下降，它取决于超级电容器的电容。在正常操作中，电容器作为辅助电源将为下一次负载变化自动充电。这给了连接到 SOFC 系统的控制器一个附加的电压，以保持电压在它的参考值内。通过使用一个电容器避免电压突然下降，该控制器可以更容易地将操作变量的电压达到目标值。这个概念随后会被描述。

假设流过电容器和燃料电池的电流分别是 I_{uc} 和 I_{fc}。然后，I_{fc} 可以根据需求电流表示为

$$I_{fc} = I_d - I_{uc} \tag{11.24}$$

从电容的定义可知，流过电容器的电流与直流电压源并联可表示为

$$I_{uc} = -C\frac{dV_s}{dt} \tag{11.25}$$

式中，C 是法拉电容；V_s 是电容器两端的端电压；DC 是直流电压源。

然后，通过燃料电池的电流可以被改写为

$$I_{fc} = I_d + C\frac{dV_s}{dt} \tag{11.26}$$

燃料电池的堆栈电压，如式（11.23）中所示，可以被写为

$$\begin{aligned} V_s &= V_0 - rI_{fc} \\ \Rightarrow V_s &= V_0 - r\left(I_d + C\frac{dV_s}{dt}\right) \\ \Rightarrow rC\frac{dV_s}{dt} &= V_0 - rI_d - V_s \end{aligned} \tag{11.27}$$

从式（11.27）中可以看出，该组电压采取一阶常微分方程的形式，而不是一个静态的非线性输出方程。因此，逐步增加的需求电流取代了突降的电压，电压下降的减慢取决于 rC，这将在 11.4.2 节被证明。

11.4 仿真结果

现在，该燃料电池系统的所有部件的模型已经建立，它们可以被用来模拟在稳态和瞬态响应。上一节中描述的是在不同的操作条件下对燃料电池系统进行仿真。在仿真中使用的 BOP 参数，见表 11.1。以下各节包含在该模型基础上的仿真中。

11.4.1 燃料电池系统仿真

燃料电池系统的不同堆栈电压和组件温度下的动态响应如图 11.4 所示。燃料电池系统的堆栈电压依赖于入口的流速和温度，电压的预测并不简单。例如，增加的需求电流与燃料电池的燃料和空气的入口温度相关联，以及氢气生产速度的增加和重整装置整齐地下降。由于局部压力不仅局限于一个流速功能而且有温度的功能，H_2 和 H_2O 的分压的分别减小和增大会导致堆栈电压的附加损耗。

表 11.1　设备（BOP）参数的平衡

空气热交换器：	
$D_{i,tube}=0.20m$，$D_{o,tube}=0.205m$，$D_{i,shell}=0.40m$，$L=10m$	
燃料热交换器：	
$D_{i,tube}=0.05m$，$D_{o,tube}=0.055m$，$D_{i,shell}=0.10m$，$L=200m$	
重整装置：	重整装置容积：
反应面积，$A_{RX}=1000m^2$；	$V_R=10m^3$
燃烧器：	
燃烧器容积，$V_B=1m^3$	

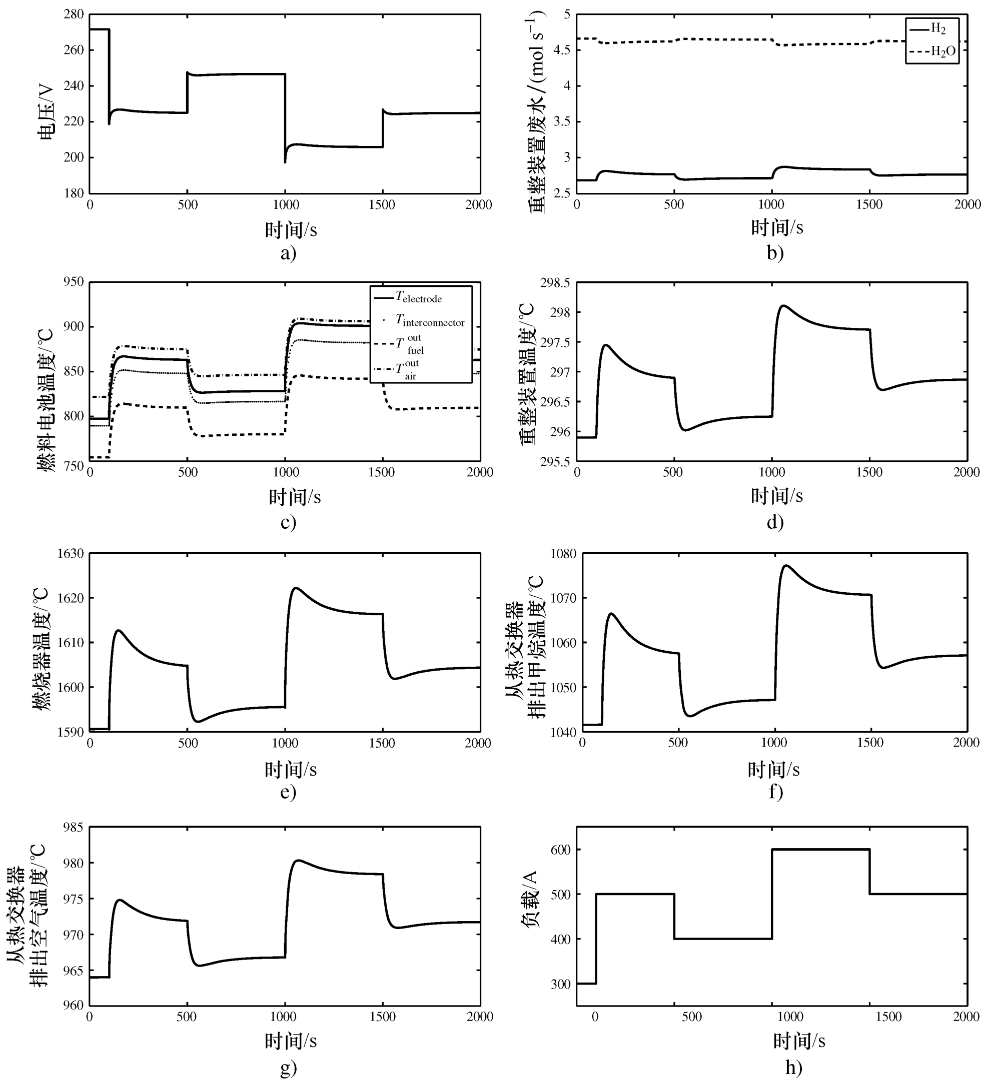


图 11.4　由于负载变化引起的燃料电池系统瞬态响应

对于增加甲烷的流速，燃料电池的堆栈电压的增加是因为氢气的增加和水蒸气的分压减小。重整装置产氢的增加，同样的重整装置的温度、燃料电池入口的温度和燃烧器的排气温度也增加。氢气流速的增加增加了局部压力，最终导致堆栈电压增加，如图 11. 5 所示。

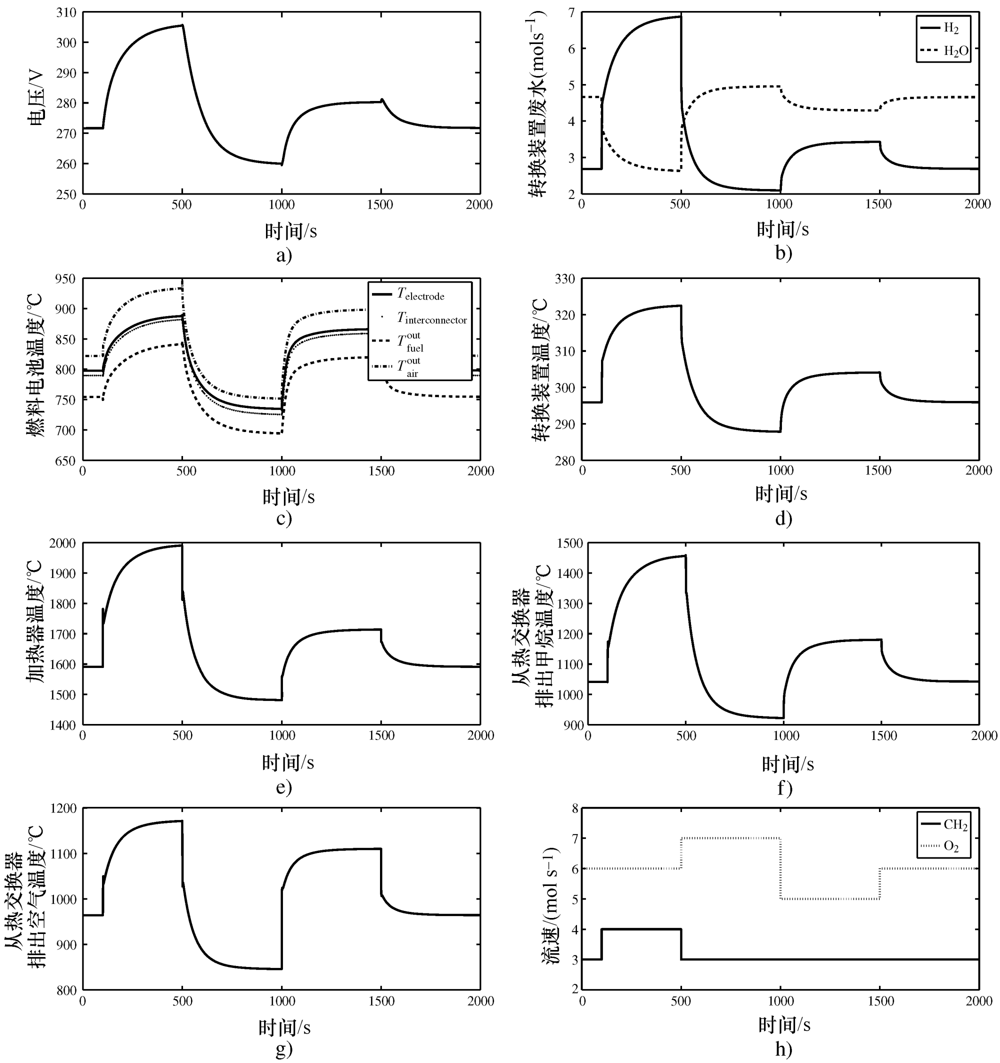


图 11. 5　由于机械流动速率和空气变化引起的燃料电池系统瞬态响应

按照相同的方式，该系统模型可以模拟不同输入工艺条件和流速的响应。因此，这些模拟可以用于经营模式发生重大变化前，进行预期调查燃料电池系统的响应。它可以用于设计和验证控制器，用于燃料电池系统，这将在 12 章进行说明。

11.4.2　超级电容器 SOFC 堆栈

对于一个独立的 SOFC，由于需求电流的改变输出电压会存在瞬间变化，如图 10.4所示。然而，当燃料电池与超级电容器并联时，合并的系统表现为一阶动态。当电容器的电容增加时，电容的负载能力也在同一时间增加，电压会急剧下降直至平滑（见图 11.6b）。换言之，电容器的工作原理类似过滤器的需求电流。最终的稳态堆栈电压在所有情况下是相同的，从式（11.27）看更明显。

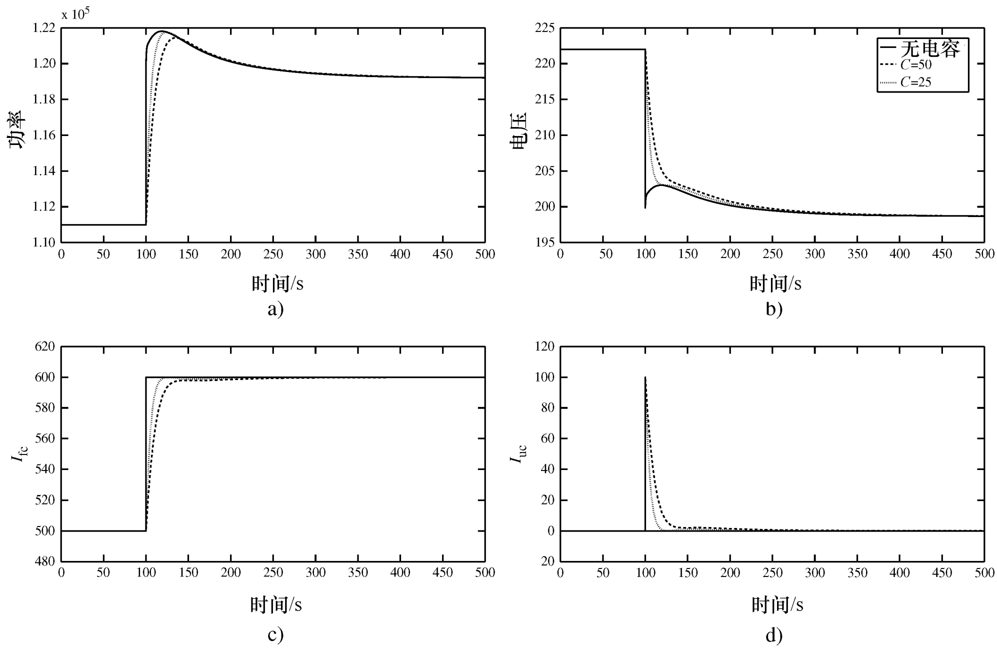


图 11.6　燃料电池与电容器动态响应

a）电源　b）电压　c）燃料电池电流负载　d）电容器的电流负载

11.5　注解和参考文献

这项工作的目的是制备具有 BOP 模型的 SOFC 模型，形成系统级控制相关的动态模型。合成并不简单的是现有模型的一个简单的组装。例如，为了使该系统可控，相关联的差分方程的附加组分是必要的，由于负载的阶跃变化要避免电压突降。此外，为了使该系统模型与控制有关，许多组件模型，包括电池模型必须重新建立。

燃料电池系统的组件的简化动态模型已经被开发出来，模拟燃料电池系统的动态行为。这些详细的推导模型可以在 Murshed 等人（2007）的结果中发现。其他实例的 SOFC 系统的工艺流程可以在 Braun（2002）的结果中发现。

第 12 章　板式 SOFC 系统的模型预测控制

12.1　简介

在前面的章节中，燃料电池系统已经建立。燃料电池系统包括很多组成部分，除了燃料电池本身外，还包括可实现高效操作这些组成部分的理想操作系统。因此，需要设计一个系统控制器并要求处理多元目标和满足设置组件的安全运行极限的限制。燃料电池的非线性动态使得这比设计一个调节控制器更具挑战性。

在操作燃料电池系统时，由于受燃料电池物理和操作的限制，并为了满足目标负载和功率要求的需要，需要满足一些约束条件。单回路和多回路的 PID 控制器，往往足以满足额定工作条件下的要求，但它们无法处理极限约束状况，在这种情况下容易发生故障。此外，模型的非线性和约束的非线性，需要考虑到控制器的设计中去。非线性模型预测控制（NMPC）是一个行之有效的线性模型预测控制（LMPC）的延伸，可以为这种应用而设计。

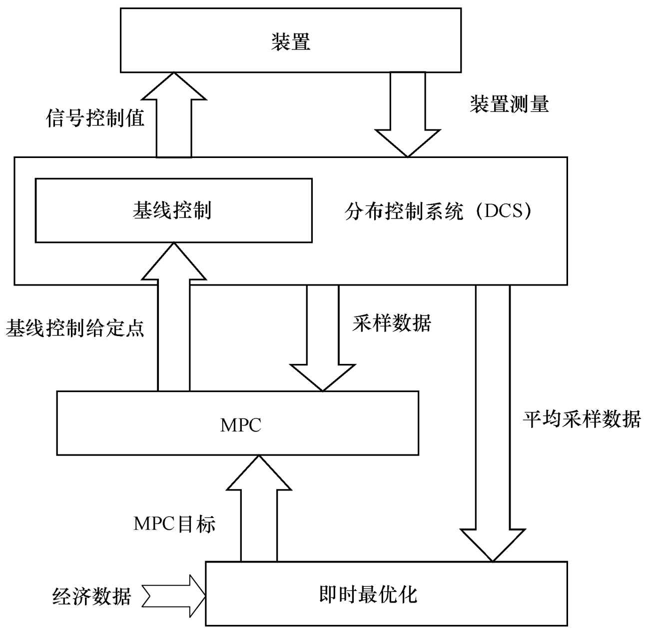


图 12.1　层级控制

由于计算的要求，通常通过诸如 LMPC 或 NMPC 等多变量的控制器计算控制动作，每分钟或更长的时间间隔计算一次。所计算的控制动作被发送到设定值的基极层控制器上，作为模型预测控制（MPC）的操作变量（MV）。像流量、压力或温度控制器等基材层控制器，大多数情况下，在控制阀上做出动作，以达到 MPC 系统的控制目标。在一般情况下，基材层控制器采用 PID 或类似的单变量控制器设计，这是当今分布式控制系统（DCS）的一个不可分割部分。根据不同的控制目标，基材层控制器可以在模型预测控制系统的内部和外部建模。在大多数情况下，它们被放置在模型预测控制系统的外面，使得这些调节器在下一个循环周期来临前正确地到达其设定点。在某些情况下，在稳态优化的高级层顶部的一些多变量控制器上实现的最佳平面的操作。图 12.1 显示了控制层级的概况。

12.2　控制目标

任何设备的主要制约因素源于其设计容量。因此，对于 SOFC，主要目标是为所需的功率提供最佳效能，最重要的是在一定的范围内运转燃料电池。SOFC 的工作温度高于其他燃料电池，由于温度分布在阳极、阴极和电解质等组件的不同造成热应力的不同，这会对燃料电池的完整性造成伤害。因此，保证电池最高温度低于某一阈值变得尤为重要，另外，阳极和阴极之间的燃料和空气的局部压力差也是值得关注的问题，想要保证压力差低于一定值，这取决于设计。除了燃料电池本身外，来自辅助设施的设计引起的燃料利用率、燃料和空气对流速度的局限也是影响燃料电池系统的重要因素之一。例如，下限值施加于蒸汽和碳比率的预重整装置的入口处，以避免碳沉积，这也是一个操作燃料电池系统的限制因素。

因此，设计一个燃料电池系统的控制器用于解决以下问题：①处理瞬态操作引起燃料电池的性质急剧变化的情况；②有效地达到目标负载；③满足不同组件安全操作极限的约束条件。第一个问题，可以通过外加的电容器处理瞬时电压变化引起负载。

另外两个问题可以通过具有处理约束能力的多变量控制器来解决，通常使用 NMPC。

操作燃料电池的控制变量（CV）的公共控制目标可以定义如下：①维持其所需的功率输出水平；②保持电池的最大温度低于临界值；③保持所有电源输出的燃料利用率在最佳工作范围内（Aguiar 等人，2005；Fue，2004；Padullés 等人，2000；Zhang 等人，2008；Zhu 和 Tomsovic，2002）。燃料和空气流速可以通过改变操作变量（MV）以达到上述目的。此外，根据不同的目标、电流密度、燃料入口温度和空气可以被定义为一个操作变量或干扰变量（DV）。对于单变量控制的情

况，可通过使用燃料比作为操作变量，可以使用的空气流量可以用于控制燃料利用率或电池温度。如果燃料利用率可以通过正在移动的空气流量控制，温度控制装置可用于保证电池温度低于其最大极限值。在这种情况下，当电池温度超过其可接受的上限时，过热控制将增加移动的空气流量。此外，为辅助温度控制设置一个预热交换器，可用于帮助调节所需的燃料和空气的入口温度。可以设计一个过压控制器，以保证阳极和阴极之间的压力差低于其设定容量。燃料比和空气利用率可作为不同压力过载控制器的操作变量。

由于控制目标是提高负荷跟踪系统的性能，堆栈电压可以被认为是一个控制变量，而燃料比，蒸汽和空气（或氧气）的流速可作为燃料电池和燃烧器的控制器的操作变量。流速可以通过使用适当的压缩机、鼓风机、阀门来控制，也可以用 ODE 模型在数学上表现。在这个特殊的情况下，废气热回收单元和热交换器分流比也可以作为一个操作变量。负载和流动的温度作为干扰系统的因素。图 12.2显示的是通过 LMPC 和 NMPC 研究控制系统性能的概略图。

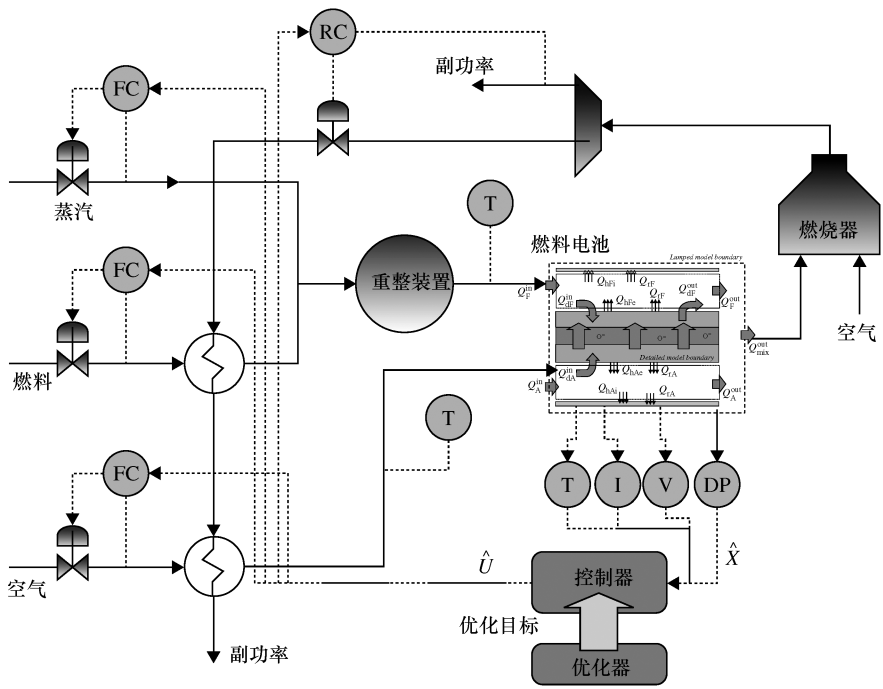


图 12.2　多变量控制和优化器的 SOFC 系统

12.3　状态估计：UKF

LMPC 可通过使用 MATLAB®的 MPC 工具箱的内置检测器评价其不可测量的状

态。模拟 NMPC 做一个合适的编码以适应燃料电池系统是不容易的，作者使用 MATLAB®模拟 SOFC 系统，并设计相应的 NMPC 系统。在实际情况下，状态是不可检测的，在这项工作中 UKF 被应用到所述燃料电池系统状态的评估中。UKF 已在第 5 章中描述。MATLAB®的 ODE 求解器被用于 UKF 生成离散数据的一个离散模型。UKF 算法测试通常是用于 NMPC 之前，用来评估单机燃料电池和整个燃料电池系统的不可测状态的表现，后面将给出仿真结果。

图 12.3 表示了由燃料电池的集总模型描述的一个独立的燃料电池的预估的局部压力，只测量燃料电池的气体出口温度（其被假定为相同的电池温度）和堆栈电压。估计的状态与真实状态比较表明，即使状态的初始条件是不完全已知的，估计的状态也可以迅速地收敛到真实状态。

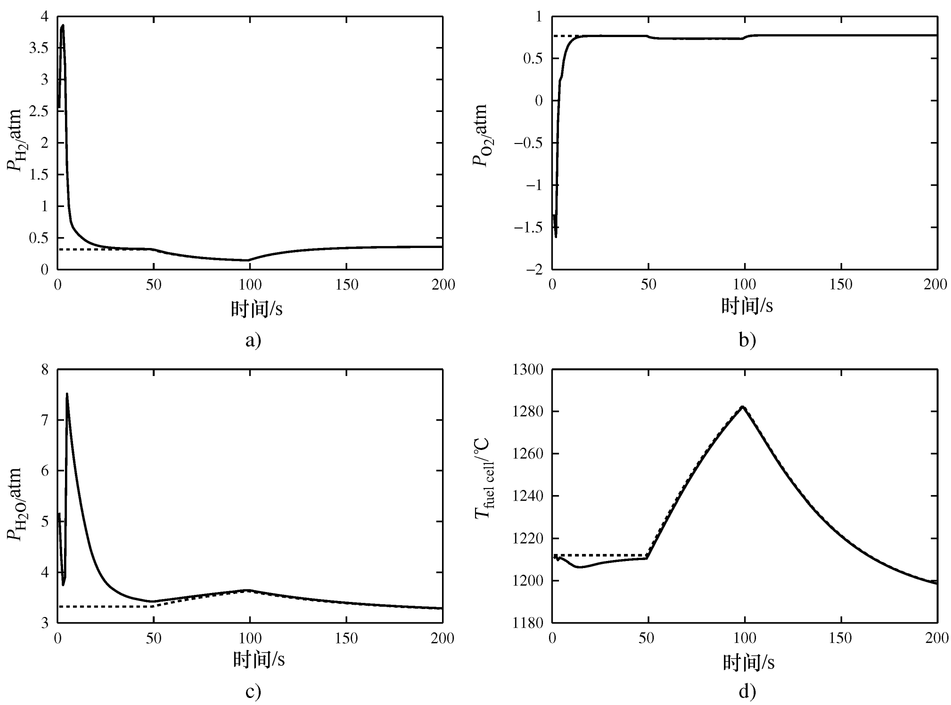


图 12.3　UKF 通过假设只有电池温度和电压是可测的条件下独立 SOFC 状态估计
（实线 = 估计状态；虚线 = 实际状态）

图 12.4 显示了 UKF 作为一个估计值用于整个燃料电池系统的表现。在这里，我们认为只有流动的温度和电池组电压是可测量的。其他燃料电池组件的温度和氢气、氧气、水蒸气的分压是估计的。很明显，从仿真估计的状态也迅速收敛到真实状态。

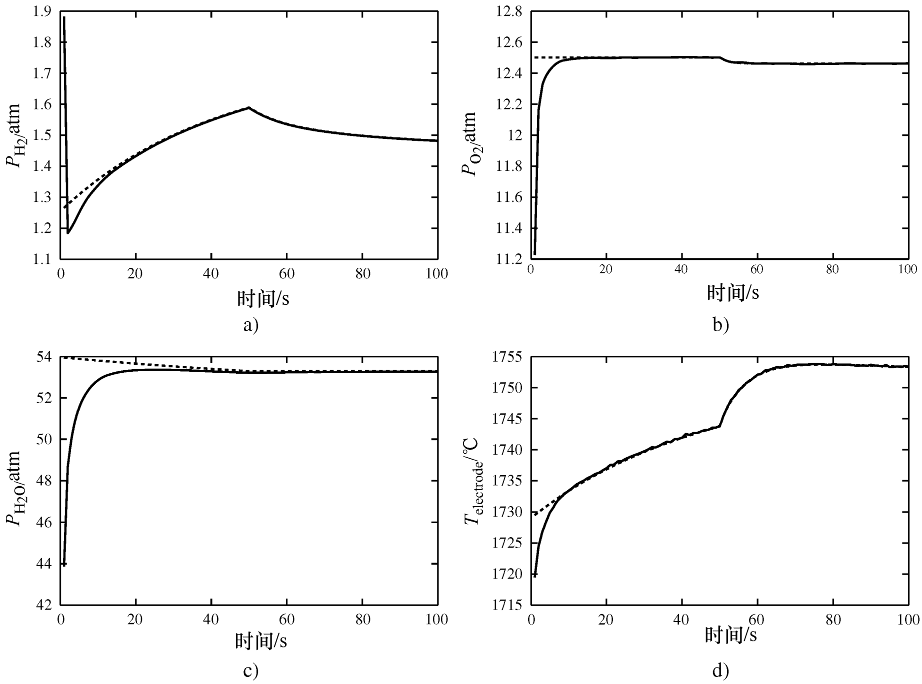


图 12.4　UKF 通过假设只有燃料流温度和堆栈电压是可测的条件下独立 SOFC 状态估计（实线 = 估计状态；虚线 = 实际状态）

12.4　稳态经济优化

燃料电池将燃料的化学能直接转化为电能，从而能有高的效率。未反应的燃料可以被燃烧产生额外的能量。燃料电池系统中的高温废气可通过使用燃气轮机，以进一步产生用于家庭和工业用途的电力或简单地产生热水和蒸汽。从燃料电池系统产生的能量可分为两部分：①直接能源；②间接能源。通过燃料电池本身所产生的能量被命名为直接能源，因为它将化学能直接转化为电能，而通过来自燃料电池未反应的燃料气体所产生的能量被命名为间接能源。

尽管间接能源为燃料电池系统提供了很大一部分的功率，但用燃料电池系统的主要目标是产生尽可能多的直接能源，尽可能地避免系统由于卡诺循环而损耗效率。换句话说，从燃料电池系统排气的能量含量越小越好。

由于排气能量主要来自燃烧反应的燃料，最大限度地减少未反应的燃料成本或热值最大化的直接能量输出是必要的。然而，未反应的燃料被烧掉，然后用预先加热系统中的燃料。通过最大限度地减少可从排气中提取的能量来替代最小化成本或未反应燃料的热值是比较合适的。这反过来相当于在指定的电压产生电力减少燃料的使用。因此，目标是使得燃料电池系统处于一个最佳的稳定状态，在一个给定的

电流和电压下，所述排气的损耗能量是最小的。在图 11.1 中显示了在燃料电池系统中经济优化的数学公式：

$$\min_{U_s} \sum \dot{n}_{i,E} \int_{T_{ref}}^{T_{AE}} C_{p,i}(T)\,dT + \dot{n}_{i,E} \frac{1-r}{r} \int_{T_{ref}}^{T_B} C_{p,i}(T)\,dT \tag{12.1}$$

$$f(X_s, U_s, W_s) = 0 \qquad [稳态模型]$$

$$V_s = V_{ref}$$

$$X_{min} < \hat{X}_s < X_{max} \qquad [NMPC 约束]$$

式中，U_s 是稳态输入（流速和分流比，r）的燃料电池系统；$\dot{n}_{i,E}$是空气热交换器的废气的流速；T_{AE}和 T_B 分别为空气热交换器和燃烧器出口温度；T_{ref}为参考温度。

以上描述的最优输入可作为处于稳定状态且排气能量最小的 NMPC 的一个参考输入。这种稳态优化器可在 MPC 的顶部被使用，在需要时可以被调用，如图 12.2所示。

直接能源最小化也可以动态地执行在每个 NMPC 步骤上，而不是如前所述的稳态优化。在这种情况下，上面的目标函数可以被合并以 NMPC 与 $\boldsymbol{Q}$、$\boldsymbol{R}$、$\boldsymbol{S}$ 和 $\boldsymbol{W}_{indirect}$间接的适当权重描述的目标函数：

$$\begin{aligned} \min_{u(t|t),u(t+1|t),\cdots,u(t+M-1|t)} J = & \sum_{i=1}^{N} [\| \hat{x}(t+i|t) - x_{ref} \|_{\boldsymbol{Q}}^2 \\ & + \| u(t+i|t) - u_{ref} \|_{\boldsymbol{R}}^2 + \| \Delta u(t+i|t) \|_{\boldsymbol{S}}^2] \\ & + \boldsymbol{W}_{indirect} \Big[\sum \dot{n}_{i,E} \int_{T_{ref}}^{T_{AE}} C_{p,i}(T)\,dT \\ & + \dot{n}_{i,E} \frac{1-r}{r} \int_{T_{ref}}^{T_B} C_{p,i}(T)\,dT \Big] \end{aligned} \tag{12.2}$$

$$\hat{x}(t+1) = f(\hat{x}(t), u(t), w(t))$$

$$y_{min} \leqslant \hat{y}_i \leqslant y_{max}$$

$$u_{min} \leqslant u_i \leqslant u_{max}$$

$$\Delta u_{min} \leqslant \Delta u_i \leqslant \Delta u_{max}$$

但是，这个公式有两个缺点：首先，计算负担太高，该系统主要在电流偶尔改变的稳定状态起作用；其次，加权矩阵 $\boldsymbol{Q}$、$\boldsymbol{R}$、$\boldsymbol{S}$ 和 $\boldsymbol{W}_{indirect}$在目标函数将确定系统是否是动态控制或稳定状态时进行了优化。另一方面，如果该方案处于离散状态，那么这两个目标可以是确定的，且不会使任何一个目标有牺牲。因此，SOFC 系统将被优化并以最低的能量在稳定状态下如式（12.1）所描述的目标函数运行。通过参考输入的 NMPC 仿真结果将在 12.5.3 节给出。

12.5 控制和仿真

在本节中，通过模拟研究模 LMPC 和 NMPC 的性能。LMPC 和 NMPC 应用于单机型

燃料电池和燃料电池系统，以及连接在与燃料电池并联的电容器。结果如下节所述。

12.5.1 LMPC

MPC 的理论已经在第 6 章中讨论过。在本节中，MATLAB®的 MPC 工具箱被用于 LMPC 燃料电池系统。在燃料电池系统中，在这种情况下，使用 MATLAB®线性系统的分析工具箱对标称工作点进行线性化。不可测状态的估计已经被集成到 MPC 公式中。因此，无外部状态的估计量是需要的。

图 12.5 显示了燃料电池系统和通过 LMPC 从 500A 到 450A 变化的较低负荷干扰下执行控制动作的瞬态响应。该 LMPC 的预测和控制的视野有 10 个和 3 个，采样间隔为 1s。在这种情况下，LMPC 能正确调节系统目标。

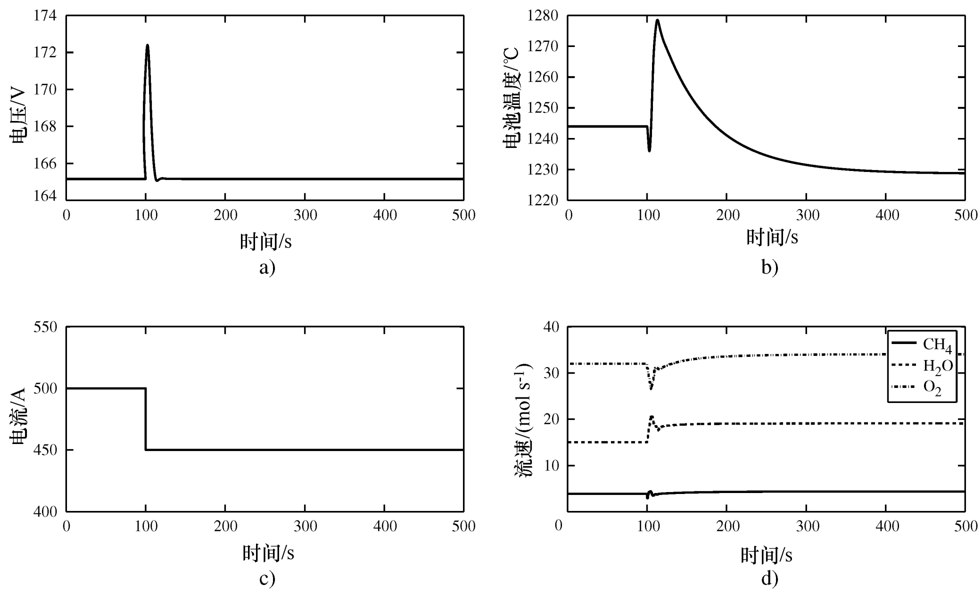


图 12.5 负载变化从 500A 到 450A 的响应，PH = 10，CH = 3，T_s = 1s

LMPC 成功地抑制住小负荷的干扰，将几种不同程度变化的负载应用到系统中，以测试不同的预测和控制。但是，可以从图 12.6 看出，LMPC 不能够通过调节电流从 500A 到 300A 来控制电压，相同的变量下最终显示的结果不稳定。

LMPC 预测范围有限，所以这种类型的不稳定性并不难预见。应该记住的是，非线性模型线性化首先围绕初始稳态运行条件再对应到 500A 负载。该线性化模型通过采取采样间隔 T_s = 1s 离散化。这种线性化和离散化模型能够很好地处理 LMPC 其标称工作状态。由于非线性的燃料电池系统模型的建立时间约 500s，但仅 10s 的预测范围的线性化模型来控制该系统的局限性太大。

LMPC 通过增加预测范围和采样的时间间隔分别为 30s 和 10s 再应用到燃料电池系统。在 300s 的较长的预测时间状况下，可以使系统更稳定，如图 12.7 所示。但根据资料发现，它需要更长的时间和更多的控制使系统在 LMPC 的控制之下。

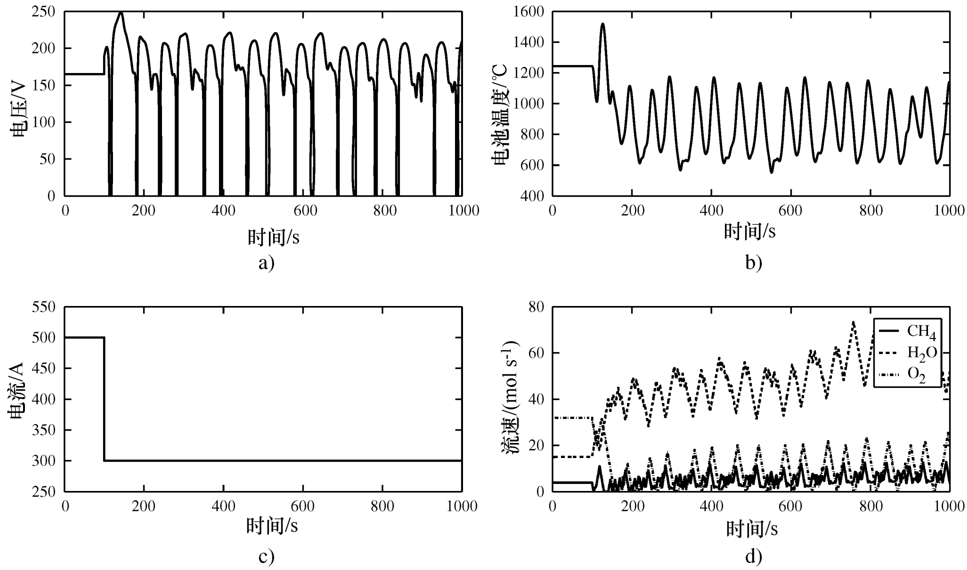


图 12.6　负载变化从 500A 到 300A 的响应，PH = 10，CH = 3，T_s = 1s

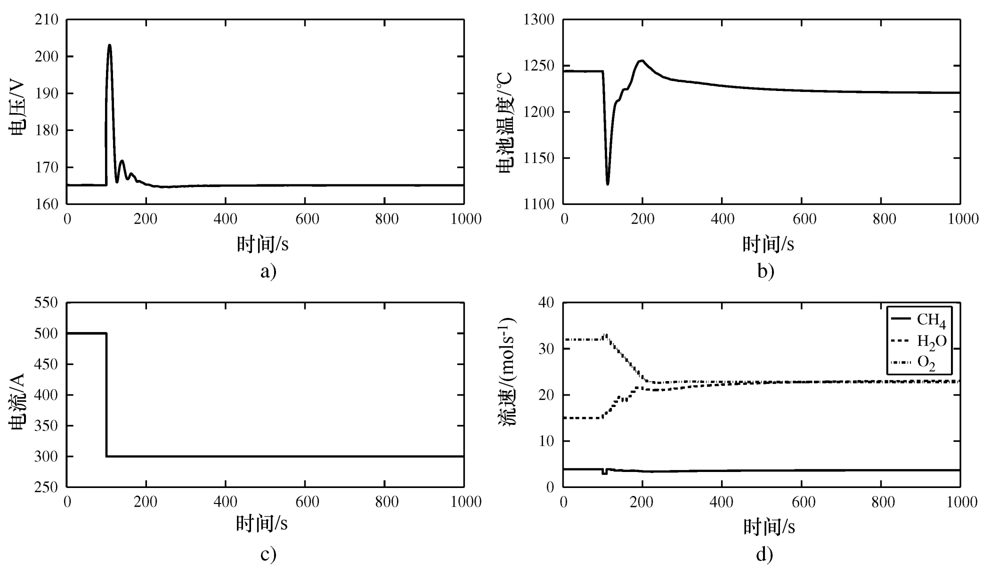


图 12.7　负载变化从 500A 到 300A 的响应，PH = 30，CH = 3，T_s = 10s

尽管 LMPC 可以稳定小负荷干扰系统，但它无法控制具有更高负载干扰的系统，除非适当地调整负荷。对于更频繁和更高的负载变化，这可能会成为一个问题，因为它需要持续关注它调整到不同的操作条件。这促使人们更积极地采用不受操作区间限制的 NMPC。NMPC 的仿真结果见 12.5.2 节。

12.5.2　NMPC

仿真也常通过在两个独立的燃料电池和燃料电池系统中应用 NMPC 来表现不同的负载引起的变化，并与 LMPC 比较。如前面提到的，非线性的连续时间模型的离散化可以通过正交方法来完成，不可测状态通过使用 UKF 估计。在所有情况下，NMPC 比 LMPC 更有优势。采用 NMPC，采样间隔 $T_s=1s$ 的预测和控制区域分别为 5 和 2，就已经足够用于所有情景的控制系统。更高的预测和控制区域需要更大的计算能力和时间，这是 NMPC 的固有缺点。

图 12.8 显示了 NMPC 连接了给定电压值的电容器的独立燃料电池的响应信息。从响应信息看，NMPC 可在很短的时间内使系统进入新的目标。

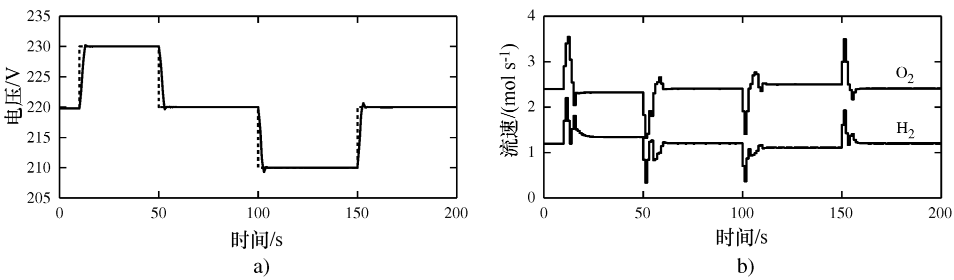


图 12.8　独立燃料电池目标电压变化的响应

图 12.9 显示了电压的响应，以及由 NMPC 针对燃料电池系统需求电流变化而采取的一系列动作。输出电压和控制操作在任何情况下都可在 10s 内做出最优反应。在这种情况下，电容器被连接在与上述燃料电池上以查看效果。已经发现，即使没有电容器，NMPC 能在相当短的系统内稳定下来，而 LMPC 需要更长的预测时间和设置时间来稳定系统。

12.5.3　优化

运行燃料电池系统以产生最大直接能源的想法已在 12.4 节研究了。该仿真将被用来找出一组针对不同载荷分布的最优输入，这可以作为在 NMPC 目标函数的参考输入。其结果见表 12.1。

表 12.1　最小间接能源的最优燃料流

电流 I/A	最优流速/(mol s^{-1})				分流比	排气能量/(kJ A^{-1})
	$\dot{n}_{CH_4}$	$\dot{n}_{O_2}$	$\dot{n}_{H_2O}$	$\dot{n}_{O_2,burner}$		
200	0.2272	0.7197	0.3327	0.2272	0.3365	1.1597
300	0.3579	1.1181	0.4165	0.3579	0.3298	1.2127
400	0.5465	1.8736	0.5111	0.5465	0.3411	1.4057
500	0.9322	4.1958	0.5863	0.9322	0.3990	2.0100
600	1.6756	14.9550	0.6690	1.6756	0.8165	4.1428

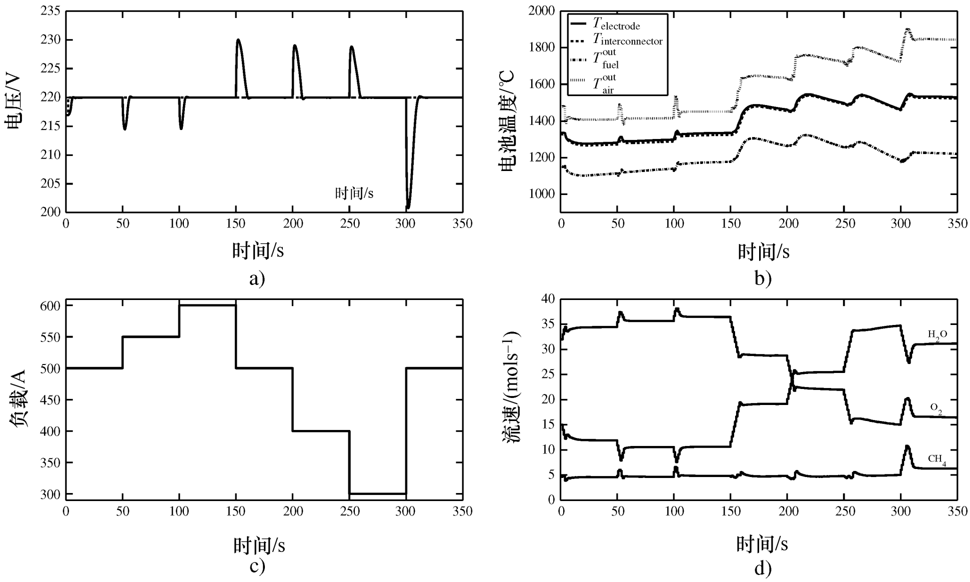


图 12.9　燃料电池系统负载变化响应

图 12.10 描述了当需求电流从 500A 变化到 700A 时，NMPC 参考输入同时发生变化，具体变化见表 12.1。该系统输出迅速达到其目标。然而，不像以前的响应，所采取的 NMPC 动作按照最佳的稳定状态设置。

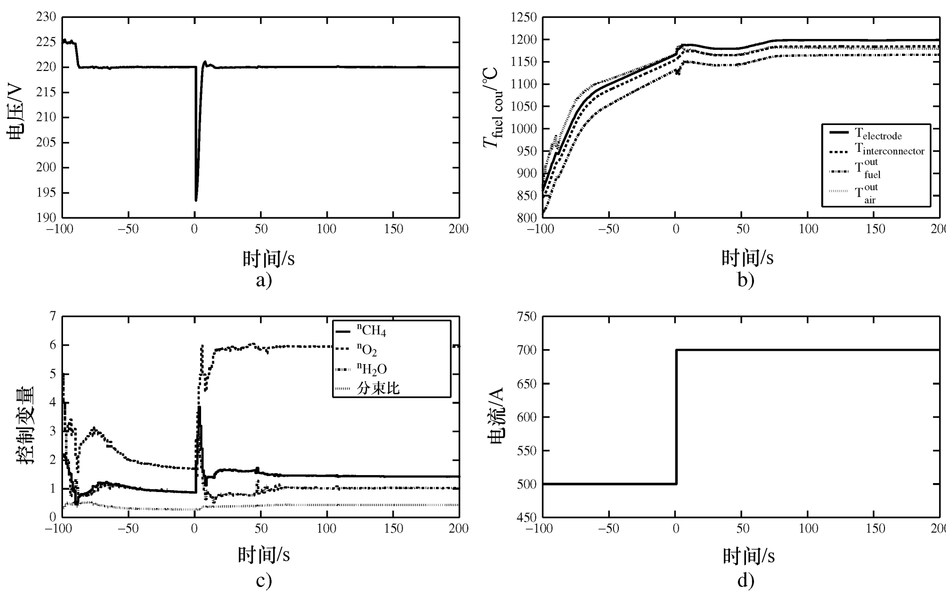


图 12.10　能量最小化的最优稳态输入目标负载变化从 500A 到 700A 的响应

从表 12.1 和图 12.10 可明显看出，在最小的间接能源操作，产生 500A 电流所

需的甲烷要比在未优化系统 5kmol s^{-1} 的流速下（见图 12.9d）的消耗量低很多。但是，也注意到，在较高电流需求条件下，燃料流率要比低电流条件下高得多，从最优化的效率看这是显而易见的，定义排气能量为每安培电流负载千焦耳。由于重整装置和热交换的限制，燃料电池系统在较高的负载下工作效率较低。这些系统组件被设计成在 200 ~ 300A 的电流负载下工作。随着负载的增加，甲烷与氢气在重整装置中的转换率降低。依次将未反应甲烷的温度和废气的能量含量增加。更高容量的重整装置将降低在较高负载条件下的效率。

12.6 结果与讨论

本章的主要目的是开发出在第 10 章和第 11 章描述过的适用于独立 SOFC 和 SOFC 系统的控制器，其中两套模型以及 BOP 和一个电容并联，公式化地运行燃料电池。

所有描述的模型是通过第一性原理的知识开发的，从而具备典型的非线性和第一性原理模型连续时间特性。LMPC 和 NMPC 已经应用到系统中，以比较它们的性能。对于 LMPC，通过使用 MATLAB 的 MPC 工具箱，因此没有明确要求输入/输出数据的不可测状态和半离散化的模型。NMPC 的实施需要离散化，这是通过使用正交方法完成的。而另一方面，不可测状态通过 UKF 估计。

从模拟中显而易见的是，简单非线性系统的 LMPC 工程在接近额定工作条件下与线性模型进行近似。在这项工作中，LMPC 适用于小负载干扰。当负荷变化达到 50A，更高的预测和控制时域是必需的，它可以简单地通过增加时间间隔来减少计算量。然而，对于更高的负载变化，LMPC 调节电压失败并最终表现出不稳定的行为。

另一方面，NMPC 可以控制较大负载干扰的系统，只要它不超过该燃料电池系统的设计容量。NMPC 的预测和控制区间可以是小的，而这又减少了计算量。

稳态优化的解决方案不是唯一的。根据不同的初始猜测，稳态化可收敛到不同的点，所有这些都满足约束条件和工艺操作条件。表 12.2 提供了在相同的目标负载和不同初始猜测值下的稳态解。

表 12.2　适用在不同初始条件下的不同稳态解

需要电流 I/A	解	最优流速/(mol s^{-1})				最优分流率
		$\dot{n}_{CH_4}$	$\dot{n}_{O_2}$	$\dot{n}_{H_2O}$	$\dot{n}_{O_2,burner}$	
200	解 1	6.0679	65.2014	8.1841	6.0679	0.5851
	解 2	0.2272	0.7197	0.3327	0.2272	0.3365
300	解 1	2.4745	22.7990	3.3505	2.4745	0.5406
	解 2	0.3579	1.1181	0.4165	0.3579	0.3298
400	解 1	1.0942	6.3157	1.4062	1.0942	0.4316
	解 2	0.5465	1.8736	0.5111	0.5465	0.3411

这不能简单地决定这些解决方案哪个是最优的，哪一个是次优的。最佳的解决方案可以选取两种方式：首先，比较目标函数的最终值和选择的解决方案是最小的；其次，设定目标的 MV 是最经济可行的，例如，燃料流率，选择具有最小燃料流量的溶液。这也可以通过引入燃料流率的罚因子来完成稳态目标函数。数学公式可以写作：

$$\min_{U_s} \sum \dot{n}_{i,E} \int_{T_{ref}}^{T_{AE}} C_{p,i}(T)\,dT + \dot{n}_{i,E} \frac{1-r}{r} \int_{T_{ref}}^{T_B} C_{p,i}(T)\,dT + \| \dot{n}_{CH_4} \|_2^{Q}$$

$$f(X_s, U_s, W_s) = 0 \qquad [\text{稳态模型}]$$

$$V_s = V_{ref}$$

$$X_{min} < \hat{X}_s < X_{max} \qquad [\text{NMPC 约束}]$$

选择最小流速有一个固有的缺点。如果负载与最小流速在运行过程中突然增加，电压会下降很多，造成电池的损坏。一种解决这个问题的方法是使用电容器作为辅助动力源，这已经深入讨论。

12.7　注解和参考文献

在过程控制的世界里，一个宽范围的线性和非线性控制技术已被开发，并在工业上应用。特别是在过去的十年中，廉价计算能力的出现，从传统 PID 控制器转移到以前难以实现的控制器，如非线性模型预测控制器的趋势正在被发掘。相同的趋势是，为燃料电池设计的控制器正被发掘。一些燃料电池系统的控制研究已经发表。大多数论文处理不同类型的燃料电池包括 PEMFC、MCFC 和 SOFC，通过采用局部或多回路控制器（Inui 等人，2006；Kandepu 等人，2006；Li 等人，2005；Stiller 等人，2006；Zhu 和 Tomsovic，2002）。不同的单变量和多变量控制策略的详细讨论可以在 Huang 等人（2011）的文章中找到。

一些研究者最近将他们的注意力集中在 SOFC 的 MPC 应用上（Murshed 等人，2010；Wang 等人，2007；Zhang 等人，2008）。在前人的工作基础上，他们开发了一个将 NMPC 应用到板式 SOFC 的闭环反馈控制策略。主要控制目标是通过操纵电流密度、燃料和空气流速来控制输出功率、燃料利用率和温度。一些作者使用模糊模型替代第一性模型来构建模型预测控制系统（SOFC）（Jurado 和 Ortega，2006；Wu 等人，2008；Yang 等人，2009）。

LMPC 或 NMPC 的实施也可能需要不可测状态的连续模型和估算离散。参见 Burns（1992）、Finlayson（1980b）、Lapidus 和 Seinfeld（1971）、Henson 和 Seborg（1997）、Stephens（1984）以及 Villadsen 和 Michelsen（1978）的文献并进一步阅读不同的离散化方法。Julier（2002）、Julier 等人（1997）、Lee（2005）、Rui 和 Chen（2005）、Haykin（2001）以及 Wan 和 Merwe（2000，2001）提供了 UKF，已在本章中就估算不可测状态做了很好的回顾。

附录　性质和参数

A.1　参数

仿真和模型参数的输入条件选择依据 Campanari 和 Iora（2004）的文献，见表 A.1和表 A.2。

表 A.1　仿真输入条件

符　　号		描　　述
R_{load}	$=4(\Omega)$	外部载荷电阻
P_{fuel}^{in}	$=1(atm)$	燃料流进入压力
T_{fuel}^{in}	$=823(K)$	燃料流进入温度
u_{fuel}^{in}	$=0.927(m\ s^{-1})$	燃料流进入速率
$\chi_{CH_4}^{in}$	$=0.173$	CH_4 输入摩尔分数
$\chi_{H_2O}^{in}$	$=0.284$	H_2O 输入摩尔分数
$\chi_{H_2}^{in}$	$=0.258$	H_2 输入摩尔分数
χ_{CO}^{in}	$=0.057$	CO 输入摩尔分数
$\chi_{CO_2}^{in}$	$=0.228$	CO_2 输入摩尔分数
P_{air}^{in}	$=1(atm)$	阴极处空气流输入压力
T_{air}^{in}	$=1104(K)$	阴极处空气流输入温度
u_{air}^{in}	$=7.79(m\ s^{-1})$	阴极处空气流输入速率
$\chi_{O_2}^{in}$	$=0.21$	O_2 输入摩尔分数
P_{inj}^{in}	$=1(atm)$	注入空气流输入压力
T_{inj}^{in}	$=1104(K)$	注入空气流输入温度
u_{inj}^{in}	$=12.08(m\ s^{-1})$	注入空气流输入速率

表 A.2　模型参数

符　　号		描　　述
几何参数		
A	$=1(\mathrm{cm}^2)$	电池面积
r_3	$=11\times10^{-3}(\mathrm{m})$	外部电池管半径
r_2	$=8.66\times10^{-3}(\mathrm{m})$	内部电池管半径
r_1	$=6\times10^{-3}(\mathrm{m})$	外部注入管半径
r_0	$=5\times10^{-3}(\mathrm{m})$	内部注入管半径
L_a	$=0.1\times10^{-3}(\mathrm{m})$	阳极厚度
L_c	$=2.21\times10^{-3}(\mathrm{m})$	阴极厚度
L_e	$=0.04\times10^{-3}(\mathrm{m})$	电解质厚度
F_{c-t}	$=0.69$	视觉因子:电池到注入管
固体材料性质		
ρ_{cell}	$=4592(\mathrm{kg\ m^{-3}})$	SOFC 电池整体密度
$C_{p,\mathrm{cell}}$	$=740(\mathrm{J(kgK)^{-1}})$	SOFC 电池整体比热
k_{cell}	$=2.0(\mathrm{W(mK)^{-1}})$	阳极整体电导率
$\varepsilon_{\mathrm{cell}}$	$=0.9$	电池内表面辐射率
ρ_{a}	$=2.98\times10^{-5}\exp\left(-\dfrac{1392}{T_{\mathrm{cell}}}\right)(\Omega\mathrm{m})$	阳极电阻率
ρ_{e}	$=2.94\times10^{-5}\exp\left(\dfrac{10350}{T_{\mathrm{cell}}}\right)(\Omega\mathrm{m})$	电解质电阻率
ρ_{c}	$=8.11\times10^{-5}\exp\left(\dfrac{600}{T_{\mathrm{cell}}}\right)(\Omega\mathrm{m})$	阴极电阻率
ρ_{tube}	$=3900(\mathrm{kg\ m^{-3}})$	注入管密度
$C_{p,\mathrm{tube}}$	$=976.8+0.2409T(\mathrm{J(kgK)^{-1}})$	注入管比热
k_{tube}	$=31.86-0.03706T+1.317\times10^{-5}T^2(\mathrm{W(mK)^{-1}})$	注入管电导率
$\varepsilon_{\mathrm{cell}}$	$=0.4$	电池内表面辐射率
重整和置换		
K_{r}	$=8542$	重整反应速率系数
K_{s}	$=100$	替换反应速率系数
E_{r}	$=95(\mathrm{kJ\ mol^{-1}})$	重整反应活化能
α	$=0.85$	重整阶数
β	$=-0.35$	重整阶数

A.2　气体性质

燃料气由多种气体组成，每一种气体都有各自的性质。组分的不同配比导致燃料气体具有不同的组成性质。因为在 SOFC 燃料通道，燃料流的摩尔分数随着重组

和置换反应的进行而变化，气体组成是时间的参数。因此，有必要知道如表 A.3 所示的每一种组成的性质，进而利用平均重量性质参数计算燃料气流的性质。

每一种组分的比热数据 C_p 参考 Chemical Engineering Handbook（Perry 等人，1997b）。在 SOFC 应用中，考虑燃料气和空气流为混合体积，因此只关注在恒定体积下的比热参数 C_v：

$$C_v = C_p - R \tag{A.1}$$

式中，R 是气体摩尔常数。

表 A.3　气体组分的固有特性参数

气　　体	摩尔质量 $M/(\mathrm{kg\ kmol^{-1}})$	临界温度 T_c/K	临界压力 P_c/Pa
CH_4	16	191	4.59×10^6
H_2O	18	647	21.94×10^6
H_2	2	33	1.32×10^6
CO	28	133	3.49×10^6
CO_2	44	304	7.39×10^6
N_2	28	126.7	3.39×10^6
O_2	32	155	5.02×10^6
空气	29	132	3.79×10^6

SOFC 的工作温度范围在 700～1500K，每一种气体组分的比热可通过一阶或二阶温度多项式进行估计，见表 A.4。估计值的最大误差也列在表中。

低压下气体黏度值根据 Reichenberg 方法进行估算（Perry 等人，1997b）：

$$\mu = \frac{AT_r}{[1 + 0.36T_r(T_r - 1)]^{1/6}} \tag{A.2}$$

表 A.4　不同气体的近似比热（C_v）（温度在 700～1500K）

气　　体	近　似　值	单　　位	最大误差（%）
CH_4	$C_{v,CH_4} = 1515 + 83.47T - 0.02053T^2$	$\mathrm{J\,(kmol\ K)^{-1}}$	0.08
H_2O	$C_{v,H_2O} = 2.075\times10^4 + 12.15T$	$\mathrm{J\,(kmol\ K)^{-1}}$	0.48
H_2	$C_{v,H_2} = 1.829\times10^4 + 3.719T$	$\mathrm{J\,(kmol\ K)^{-1}}$	0.9
CO	$C_{v,CO} = 1.966\times10^4 + 5.019T$	$\mathrm{J\,(kmol\ K)^{-1}}$	1.35
CO_2	$C_{v,CO_2} = 23690 + 31.01T - 8.875\times10^{-3}T^2$	$\mathrm{J\,(kmol\ K)^{-1}}$	0.16
N_2	$C_{v,N_2} = 1.910\times10^4 + 5.126T$	$\mathrm{J\,(kmol\ K)^{-1}}$	1.11
O_2	$C_{v,H_2} = 2.201\times10^4 + 4.936T$	$\mathrm{J\,(kmol\ K)^{-1}}$	1.49
空气	$C_{v,air} = 1.948\times10^4 + 4.936T$	$\mathrm{J\,(kmol\ K)^{-1}}$	1.12

对于无机气体：

$$A = 1.6104\times10^{-10}\left[\frac{M^{1/2}P_c^{2/3}}{T_c^{1/6}}\right] \tag{A.3}$$

式中，$T_r = \frac{T}{T_c}$、$P_r = \frac{P}{P_c}$是对比温度和对比压力；T_c 和 P_c 是临界温度和临界压力；M 为摩尔质量，见表 $A.3$。

在 700 ~ 1500K 范围内，气体黏度由一阶温度多项式进行估算，见表 A.5。

表 A.5　气体温度在 700 ~ 1500K 时的近似黏度

气　体	近　似　值	单　位	最大误差（%）
CH_4	$\mu_{CH_4} = 9.177 \times 10^{-6} + 1.755 \times 10^{-8} T$	Pa · s	1.56
H_2O	$\mu_{H_2O} = 4.567 \times 10^{-6} + 2.209 \times 10^{-8} T$	Pa · s	2.22
H_2	$\mu_{H_2} = 6.162 \times 10^{-6} + 1.145 \times 10^{-8} T$	Pa · s	1.26
CO	$\mu_{CO} = 1.399 \times 10^{-5} + 2.582 \times 10^{-8} T$	Pa · s	1.39
CO_2	$\mu_{CO_2} = 1.273 \times 10^{-5} + 2.822 \times 10^{-8} T$	Pa · s	1.98
N_2	$\mu_{N_2} = 1.435 \times 10^{-5} + 2.642 \times 10^{-8} T$	Pa · s	1.36
O_2	$\mu_{H_2} = 1.668 \times 10^{-5} + 3.108 \times 10^{-8} T$	Pa · s	1.45
空气	$\mu_{air} = 1.514 \times 10^{-5} + 2.793 \times 10^{-8} T$	Pa · s	1.39

一般来说，一种气体的热导率可通过下式同时进行估算：

$$k = \frac{\mu}{M}\left(1.30C_v + 14644.0 - \frac{2928.8}{T_r}\right) \tag{A.4}$$

甲烷的热导率在很大程度上依赖于温度，通常根据下式同时进行估算：

$$k = 10^{-7}(14.52T_r - 5.14)^{2/3}\left(\frac{C_p}{\lambda}\right) \tag{A.5}$$

式中

$$\lambda = T_c^{1/6} M^{1/2}\left(\frac{101325}{P_c}\right)^{2/3} \tag{A.6}$$

在 700 ~ 1500K，热导率由一阶温度估算，见表 A.6。

表 A.6　温度在 700 ~ 1500K 的气体近似热导率

气　体	近　似　值	单　位	最大误差（%）
CH_4	$k_{CH_4} = -0.04446 + 2.093 \times 10^{-4} T$	$W(mK)^{-1}$	0.41
H_2O	$k_{H_2O} = -0.01450 + 9.782 \times 10^{-5} T$	$W(mK)^{-1}$	0.61
H_2	$k_{H_2} = 0.08525 + 2.964 \times 10^{-4} T$	$W(mK)^{-1}$	0.22
CO	$k_{CO} = 0.01275 + 5.384 \times 10^{-5} T$	$W(mK)^{-1}$	1.42
CO_2	$k_{CO_2} = 0.005485 + 6.272 \times 10^{-5} T$	$W(mK)^{-1}$	2.36
N_2	$k_{N_2} = 0.01258 + 5.444 \times 10^{-5} T$	$W(mK)^{-1}$	1.18
O_2	$k_{H_2} = 0.01569 + 5.690 \times 10^{-4} T$	$W(mK)^{-1}$	0.79
空气	$k_{air} = 0.01329 + 5.539 \times 10^{-5} T$	$W(mK)^{-1}$	1.25

不同温度时的生成焓由下式进行计算：

$$H|_T = H|_{298.15} + \int_{298.15}^{T} C_p \mathrm{d}T \tag{A.7}$$

式中，$H_{H_2}|_{298.15} = 0(\mathrm{J\ mol^{-1}})$；$H_{O_2}|_{298.15} = 0(\mathrm{J\ mol^{-1}})$；$H_{H_2O}|_{298.15} = -241827(\mathrm{J\ mol^{-1}})$。

生成焓的估算值见表 A.7。

表 A.7　温度在 700～1500K 的气体近似生成焓

气　体	近　似　值	单　位	最大误差（%）
CH_4	$H_{CH_4} = -10.081 \times 10^4 + 62.12T$	$\mathrm{J\ mol^{-1}}$	17.3
H_2O	$H_{H_2O} = -25.790 \times 10^4 + 42.47T$	$\mathrm{J\ mol^{-1}}$	0.26
H_2	$H_{H_2} = -0.9959 \times 10^4 + 30.73T$	$\mathrm{J\ mol^{-1}}$	1.48
CO	$H_{CO} = -12.150 \times 10^4 + 31.81T$	$\mathrm{J\ mol^{-1}}$	0.37
CO_2	$H_{CO_2} = -41.700 \times 10^4 + 50.89T$	$\mathrm{J\ mol^{-1}}$	0.25
N_2	$H_{N_2} = -1.059 \times 10^4 + 31.40T$	$\mathrm{J\ mol^{-1}}$	1.73
O_2	$H_{O_2} = -1.2290 \times 10^4 + 35.12T$	$\mathrm{J\ mol^{-1}}$	2.17
空气	$H_{air} = -1.0947 \times 10^4 + 32.50T$	$\mathrm{J\ mol^{-1}}$	2.43

参考文献

Achenbach, E. (1994) Three-dimensional and time dependent simulation of a planar solid oxide fuel cell stack. *Journal of Power Sources*, **49**, 333–348.

Achenbach, E. (1995) Response of a solid oxide fuel cell to load change. *Journal of Power Sources*, **57**, 105–109.

Aguiar, P., Adjiman, C. and Brandon, N. (2005) Anode-supported intermediate-temperature direct internal reforming solid oxide fuel cell: II. model-based dynamic performance and control. *Journal of Power Sources*, **147** (1-2), 136–147.

Ahmed, K. and Foger, K. (2000) Kinetics of internal steam reforming of methane on Ni/YSZ-based anodes for solid oxide fuel cells. *Catalysis Today*, **63**, 479–487.

Allgower, F., Findeisen, R. and Nagy, Z. (2004) Nonlinear model predictive control: from theory to application. *Journal of the Chinese Institute of Chemical Engineers*, **35** (3), 299–315.

Astrom, K. and Wittenmark, B. (1990) *Computer-controlled Systems, Theory and Design* 2nd edn, Prentice-Hall.

Bergman, N. (1999) *Recursive bayesian estimation: navigation and tracking applications*. PhD thesis. Linkoping University.

Bessette, N.F. (1994) *Modeling and Simulation for solid oxide fuel cell power systems*. PhD thesis. Georgia Institute of Technology.

Bird, R.B., Stewart, W.E. and Lightfoot, E.N. (1960) *Transport Phenomena*, John Wiley & Sons, Inc., New York.

Blackburn, J., Reethof, G. and Shearer, J. (1960) *Fluid Power Control*, The Technology Press of M.I.T. and John Wiley & Sons, Inc., New York and London.

Bove, R., Lunghi, P. and Sammes, N.M. (2005) SOFC mathematic model for systems simulations – Part 2: definition of an analytical model. *International Journal of Hydrogen Energy*, **30**, 189–200.

Box, G. and Jenkins, G. (1976) *Time Series Analysis Forecasting and Control*, Holden-Day.

Braun, R.J. (2002) *Optimal design and operation of solid oxide fuel cell systems for small-scale stationary applications*. PhD thesis. University of Wisconsin-Madison.

Burns, J.A. (1992) Optimal sensor location for robust control of distributed parameter systems. *Proceedings of 33rd IEEE conference on Decision and Control*, Lake Buena Vista, FL. pp. 3967–3972.

Camacho, C. and Bordons, C. (1999) *Model Predictive Control*, Springer-Verlog London Limited.

Campanari, S. and Iora, P. (2004) Definition and sensitivity analysis of a finite volume SOFC model for a tubular cell geometry. *Journal of Power Sources*, **132**, 113–126.

Cengel, Y.A. and Boles, M.A. (1994) *Thermodynamics: An Engineering Approach*, 2nd edn, McGraw-Hill, Inc., New York.

Chan, S.H., Khor, K.A. and Xia, Z.T. (2001) A complete polarization model of a solid oxide fuel cell and its sensitivity to the change of cell component thickness. *Journal of Power Sources*, **93**, 130–140.

Chen, W., Bakshi, B., Goel, P. and Ungarala, S. (2004) Bayesian estimation via sequential monte carlo sampling: unconstrained nonlinear dynamic systems. *Industrial & Engineering Chemistry Research*, **43**, 4012–4025.

Daneshpour, N. (2009) *Control loop performance assessment with closed-loop subspace identification*. Master's thesis. University of Alberta.

Doucet, A., de Freitas, N. and Gordon, N. (2001) *Sequential Monte Carlo Methods in Practice*, Springer Verlag.

Evensen, G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99**, 10,143–10,1624.

Evensen, G. (2009) The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, **29**, 82–104.

Felder, R.M. and Rousseau, R.W. 1986 *Elementary Principles of Chemical Processes*, 2nd edn, John Wiley & Sons, Inc.

Finlayson, B.A. (1980) *Nonlinear Analysis in Chemical Engineering*, McGraw-Hill Chemical Engineering Series, McGraw-Hill International Book Co.

Fogler, H.S. (1999) *Elements of Chemical Reaction Engineering*, 3rd edn, Prentice Hall PTR, Upper Saddle River, NJ.

EG&G Technical Services (2004), Fuel Cell Handbook, Edition 7, US Department of Energy, 2004.

Garcia, C. and Morari, M. (1982) Internal model control-1. an unifying review and some new results. *Industrial and Engineering Chemistry*, **21**, 308–323.
Garcia, C. and Morari, M. (1985) Internal model control-2. design procedure for multivariable systems. *Industrial and Engineering Chemistry*, **24**, 472–484.
Garcia, C., Pratt, D. and Morari, M. (1989) Model predictive control: theory and practice – a survey. *Automatica*, **25** (3), 335.
Gemmen, R.S. and Johnson, C.D. (2005) Effect of load transients on SOFC operation – current reversal on loss of load. *Journal of Power Sources*, **144**, 152–164.
Gordon, N., Salmond, D. and Smith, A. (1993) A novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings on Radar and Signal Processing*, **140**, 107–113.
Gustafsson, F. and Hriljac, P. (2003) Particle filters for system identification of state-space models linear in either parameters or states. *Proceedings 13th IFAC Symposium on System Identification*, Rotterdam, Netherlands, pp. 1287–1292.
Hall, D.J. (1997) *Transient modeling and simulation of a solid oxide fuel cell*. PhD thesis. University of Pittsburgh.
Hall, D.J. and Colclaser, R.G. (1999) Transient modeling and simulation of a tubular solid oxide fuel cell. *IEEE Transactions on Energy Conversion*, **14** (3), 749–753.
Huang, B. and Forbes, F. (1997-2008) *Lecture Notes for CHE576: Intermediate Process Control*, University of Alberta.
Huang, B. and Kadali, R. (2008) *Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-driven Subspace Approach*, Springer-Verlag, London.
Huang, B. and Wang, Q.G. (2006) *Overview of emerging Bayesian approach to nonlinear model identification, International Workshop on Solving Industrial Control and Optimization Problems*, Gramado, Brazil.
Huang, B., Qi, Y. and Murshed, M. (2011) Solid oxide fuel cell: perspective of dynamic modeling and control. *Journal of Process Control*, **21**, 1426–1437.
Huang, B. and Xing, J. (2006) Dynamic modeling and prediction of cytotoxicity on microelectronic cell sensor array. *Canadian Journal of Chemical Engineering*, **86**, 393–405.
Incropera, F.P., Dewitt, D.P., Bergman, T.L. and Lavine, A.S. (2006) *Fundamentals of Heat and Mass Transfer*, 6th edn, John Wiley & Sons, Inc., New York.
Inui, Y., Ito, N., Nakajima, T. and Urata, A. (2006) Analytical investigation on cell temperature control method of planer solid oxide fuel cell. *Energy Conversion and Management*, **47**, 2319–2328.
Iora, P., Aguiar, P., Adjiman, C.S. and Brandon, N.P. (2005) Comparison of two IT DIR-SOFC models: Impact of variable thermodynamic, physical and flow properties. Steady-state and dynamic analysis. *Chemical Engineering Science*, **60**, 2963–2975.
Jazwinski, A. (1970) *Stochastic Processes and Filtering Theory*, Academic, San Diego, CA.
John, J.E.A. and Haberman, W.L. (1988) *Introduction to Fluid Mechanics*, 3rd edn, Prentice Hall, London, Sydney, Toronto, Mexico City, Tokyo, Rio de Janeiro.
Julier, S. (2002) The scaled unscented transformation. *Proceedings of the American Control Conference*, vol. 6, pp. 4555–4559.
Julier, S. and Uhlmann, J. (1997) New extension of the Kalman filter to nonlinear systems. *Signal Processing*, **3068**, 182–193.
Julier, S. and Uhlmann, J. (2004) Unscented filtering and nonlinear estimation. *IEEE Review*, **92**, 401–422.
Julier, S., Uhlmann, J. and Durrant-Whyte, H. (1997) A new approach for filtering nonlinear systems. *Proceedings of the American Control Conference*, pp. 1628–1632.
Jurado, F. and Ortega, M. (2006) Model based predictive control of fuel cells. *Electric Power Components & Systems*, **34** (5), 587–602.
Kalman, R. (1960) A new approach to linear filtering and prediction problems. *Transactions of the ASME. Series D, Journal of Basic Engineering*, **82**, 35–45.
Kandepu, R., Imsland, L., Stiller, C., Foss, B.A. and Kariwala, V. (2006) Control-relevant modeling and simulation of a SOFC-GT hybrid system. *Modeling Identification and Control*, **27** (3), 143.
Katayama, T. (2005) *Subspace Methods for System Identification*, Springer.
Khatibisepehr, S., Gonzalez, R., Shardt, Y. and Huang, B. (2009-2011) *Assignment solutions for CHE662 – process identification*, University of Alberta.
Knudsen, T. (2001) Consistency analysis of subspace identification methods based on linear regression approach. *Automatica*, **37**, 81–89.
Kouvaritakis, B., Rossiter, J.A. and Schuurmans, J. (2000) Efficient robust predictive control. *IEEE Transactions on Automatic Control*, **45** (8), 1545–1549.
Kwakernaak, H. and Sivan, R. (1972) *Linear Optimal Control System*, John Wiley & Sons, Inc.

Lapidus, L. and Seinfeld, J.H. (1971) *Numerical Solution of Ordinary Differential Equations*, *Mathematics in Science and Engineering*, vol. 74, Academic Press.

Larminie, J. and Dicks, A. (2003) *Fuel Cell Systems Explained*, 3rd edn, John Wiley & Sons, Inc., Chichester, West Sussex.

Lee, D.J. (2005) *Nonlinear Bayesian filtering with applications to estimation and navigation*. PhD thesis. Texas A & M University.

Lee, J. and Yu, Z. (1994) Tuning of model predictive controllers for robust perfomance. *Computers & Chemical Engineering*, **18** (1), 15.

Levenspiel, O. (1999) *Chemical Reaction Engineering*, 3rd edn, John Wiley & Sons, Inc., New York.

Li, Y., Choi, S. and Rajakaruna, S. (2005) An analysis of the control and operation of a solid oxide fuel cell power plant in an isolated system. *IEEE Transactions on Energy Conversion*, **20** (2), 381–387.

Lienhard, J.H. IV and Lienhard, J.H. V (2005) *A Heat Transfer Textbook*, 3rd edn, Phlogiston Press, Cambridge, MA.

Lindskog, P. (1996) *Methods, algorithms and tools for system identification based on prior knowledge*. PhD thesis. Linkoping University.

Ljung, L. (1999) *System Identification*, 2nd edn, Prentice-Hall.

Macdonald, J. (1987) *Impedance Spectroscopy: Emphasizing Solid Materials and Systems*, John Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore.

Meadows, E.S. and Rawlings, J.B. (1997) Model predictive control, in *Nonlinear Process Controls* (eds M.A. Henson and D.E. Seborg), Prentice Hall, Chapter 5, pp. 233–310.

Mitterdorfer, A. and Gauckler, L.J. (1999) Reaction kinetics of the Pt, $O_2(g)|c - ZrO_2$ system: precursor mediated adsorption. *Solid State Ionics*, **117**, 203–217.

Morari, M. and Lee, J. (1999) Model predictive control: past, present and future. *Computers & Chemical Engineering*, **23** (4/5), 667–682.

Morari, M. and Zafiriou, E. (1989) *Robust Process Control*, Prentice Hall.

Munson, B.R., Young, D.F., Okiishi, T.H. and Huebsch, W.W. (2009) *Fundamentals of Fluid Mechanics*, 6th edn, John Wiley & Sons, Inc., New York.

Murshed, A.M., Huang, B. and Nandakumar, K. (2003) Control of chaos in a convective loop system. *Nonlinear Dynamics and Systems Theory*, **3** (2), 203–226.

Murshed, A.M., Huang, B. and Nandakumar, K. (2007) Control relevant modeling of planer solid oxide fuel cell system. *Journal of Power Sources*, **163**, 830–845.

Murshed, A., Huang, B. and Nandakumar, K. (2010) Estimation and control of solid oxide fuel cell system. *Computers and Chemical Engineering*, **34** (1), 96–111.

Muske, K. and Edgar, T. (1997) Nonlinear state estimation, in *Nonlinear Process Control* (eds M. Henson and D. Seborg), Prentice Hall, New Jersey, Chapter 6, pp. 311–370.

Nelson, A. (2000) *Nonlinear estimation and modeling of noisy time-series by Dual Kalman filtering methods*. PhD thesis. Linkoping University.

Norgaard, M., Poulsen, M. and Ravn, O. (2000) New developments in state estimation for nonlinear systems. *Automatica*, **36**, 1627–1638.

Nowak, M. (2001) *Virus Dynamics: Mathematical Principles of Immunology and Virology*, Oxford University Press.

Oshima, M., Hashimoto, I., Takamatsu, T. and Ohno, H. (1991) Robust stability of model predictive control. *Industrial Chemical Engineering*, **31** (1), 119.

Overschee, P.V. and Moor, B.D. (1996) *Subspace Identification for Linear Systems: Theory, Implementation, Applications*, Kluwer Academic Publishers.

Ozisik, M.N. (1985) *Heat Transfer: A Basic Approach*, McGraw-Hill, Inc., New York.

Padullés, J., Ault, G. and McDonald, J. (2000) Integrated SOFC plant dynamic model for power systems simulation. *Journal of Power Sources*, **86** (1), 495–500.

Pearson, R. and Ogunnaike, B. (1997) Nonlinear process identification, in *Nonlinear Process Control* (eds M. Henson and D. Seborg), Prentice Hall, New Jersey, Chapter 2, pp. 11–110.

Perry, R.H., Green, D.W. and Maloney, J.O. (eds) (1997) *Perry's Chemical Engineers' Handbook*, 7th edn, McGraw-Hill, Inc., New York, London, Mexico City, Milan, Montreal, New Delhi, Singapore, Sydney, Tokyo, Toronto.

Qi, Y., Huang, B. and Chuang, K. (2005) Dynamic modeling of solid oxide fuel cell: the effect of diffusion and inherent impedance. *Journal of Power Sources*, **150**, 32–47.

Qi, Y., Huang, B. and Luo, J. (2006) Dynamic modeling of a finite volume of solid oxide fuel cell: the effect of transport dynamics. *Chemical Engineering Science*, **61**, 6057–6076.

Qi, Y., Huang, B. and Luo, J. (2008) 1-D dynamic modeling of SOFC with analytical solution for reacting gas-flow problem. *AIChE Journal*, **54** (6), 1537–1553.

Rao, S.S. (2009) *Engineering Optimization: Theory and Pratice*, 4th edn, John Wiley & Sons, Inc., Hoboken, NJ.

Rao, C.V. and Rawlings, J.B. (2000) Nonlinear moving horizon state estimation. *Progress in Systems and Control Theory*, **26**, 45–69.

Rawlings, J. (2000) Tutorial overview of model predictive control. *IEEE Control Systems Magazine*, **20**, 38–52.

Ristic, B., Arulampalam, S. and Gordon, N. (2004) *Beyond the Kalman Filter*, Artech House, Boston, MA.

Rui, Y. and Chen, Y. (2005) Better proposal distributions: object tracking using unscented particle filter. http://research.microsoft.com/en-us/um/people/yongrui/ps/cvpr01a.pdf, Collaboration and Multimedia Systems Group, Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399.

Seborg, D., Edgar, T. and Mellichamp, D. (1989) *Process Dynamics and Control*, John Wiley & Sons, Inc.

Sedghisigarchi, K. and Feliachi, A. (2004) Dynamic and transient analysis of power distribution systems with fuel cells – part i: fuel-cell dynamic model. *IEEE Transactions on Energy Conversion*, **19** (2), 423–428.

Shao, X., Huang, B., Lee, J., Xu, F. and Espejoi, A. (2011) Bayesian methods for multirate data synthesis and model calibration. *AIChE Journal*, **57**, 1514–1525.

Shridhar, R. and Cooper, D. (1998) A tuning strategy for unconstrained multivariable model predictive control. *Industrial & Engineering Chemistry Research*, **37**, 4003–4016.

Sitz, A., Schwarz, U., Kurths, J. and Voss, H. (2002) Estimation of parameters and unobserved components for nonlinear systems from noisy time series. *Physical Review*, **66**, 016210: 1–9.

Smith, J.M. and Ness, H.C.V. (1987) *Introduction to Chemical Engineering Thermodynamics*, 4th edn, McGraw-Hill, Inc., New York.

Söderström, T. and Stoica, P. (1989) *System Identification*, Prentice Hall International, UK.

Solheim, A. (1992) *In SOFC micromodelling, an international energy agency SOFC task report, Swiss Federal Office of Energy Brene*, Switzerland, p. 9.

Stephens, A.B. (1984) A finite difference galerkin formulation for the incompressible Navier-Stokes equation. *Journal of Computational Physics*, **53**, 152–172.

Stiller, C., Thorud, B., Bolland, O., Kandepu, R. and Imsland, L. (2006) Control strategy for a solid oxide fuel cell and gas turbine hybrid system. *Journal of Power Sources*, **158** (1), 303–315.

Treybal, R.E. (1980) *Mass-transfer Operations*, 3rd edn, McGraw-Hill, Inc., Auckland.

Tsai, T. and Barnett, S. (1997) Effect of LSM-YSZ cathode on thin-electrolyte solid oxide fuel cell performance. *Solid State Ionics*, **93**, 207–217.

Ulrich, G.D. (1993) *A Guide to Chemical Engineering Reactor Design and Kinetics*, Ulrich Research & Consulting, Inc., New Hampshire.

van der Merwe, R. (2004) *Sigma-Point Kalman filters for probability inference in dynamic state-space models*. PhD thesis. Oregon Health and Science University.

Villadsen, J. and Michelsen, M.L. (1978) *Solution of Differential Equation Models by Polynomial Approximation*, Prentice Hall Inc., USA.

Wagner, N., Schnurnberger, W., Müller, B. and Lang, M. (1998) Electrochemical impedance spectra of solid-oxide fuel cells and polymer membrane fuel cells. *Electrochimica Acta*, **43**, 3785–3793.

Wan, E. and van der Merwe, R. (2000) The unscented Kalman filter for nonlinear estimation. *Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control*, IEEE, Lake Louise, Alberta, Canada, http://www.ene.unb.br/~gaborges/disciplinas/pe/papers/wan2000.pdf.

Wan, E.A. and van der Merwe, R. (2001) The unscented Kalman filter for nonlinear state estimation, in *Kalman Filtering and Neural Networks* (ed. S. Haykin), John Wiley & Sons, Inc., Chapter 7.

Wan, E., van der Merwe, R. and Nelson, A. (2000) Dual estimation and the unscented transformation, in *Neural Information Processing Systems 12* (eds S. Solla, T. Leen and K. Mïller), MIT Press.

Wang, S. and Huang, B. (2011) System identification of tubular SOFC. *Technical Report 2011-05*, University of Alberta.

Wang, X., Huang, B. and Chen, T. (2007) Data-driven predictive control for solid oxide fuel cells. *Journal of Process Control*, **17** (2), 103–114.

Welty, J.R., Wicks, C.E. and Wilson, R.E. (1984) *Fundamentals of Momentum, Heat, and Mass Transfer*, John Wiley & Sons, Inc., New York.

Wu, X.J., Zhu, X.J., Cao, G.Y. and Tu, H.Y. (2008) Predictive control of SOFC based on a ga-rbf neural network model. *Journal of Power Sources*, **179** (1), 232–239.

Xue, X., Tang, J., Sammes, N. and Du, Y. (2005) Dynamic modeling of single tubular SOFC combining heat/mass transfer and electrochemical reaction effects. *Journal of Power Sources*, **142**, 211–222.

Yang, J., Li, X., Mou, H.G. and Jian, L. (2009) Predictive control of solid oxide fuel cell based on an improved takagi-sugeno fuzzy model. *Journal of Power Sources*, **193** (2), 699–705.

Zhang, X., Chan, S., Ho, H. *et al*. (2008) Nonlinear model predictive control based on the moving horizon state estimation for the solid oxide fuel cell. *International Journal of Hydrogen Energy*, **33** (9), 2355–2366.

Zhu, Y. and Tomsovic, K. (2002) Development of models for analyzing the load-following pwerformance of microturbines and fuel cells. *Electric Power Systems Research*, **62**, 1–11.

Copyright © 2013 John Wiley & Sons, Ltd.

All Right Reserved. This translation published under license. Authorized translation from English language edtion, entitled < Dynamic Modelling and Predictive Control in Solid Oxide Fuel Cells: First Principle and Data-based Approaches >, ISBN: 978-0-470-97391-2, by Biao Huang, Yutong Qi, AKM Monjur Murshed, Published by John Wiley & Sons. No part of this book may be reproduced in any form without the written permission of the original copyrights holder.

本书中文简体字版由机械工业出版社出版，未经出版者书面允许，本书的任何部分不得以任何方式复制或抄袭。版权所有，翻印必究。

北京市版权局著作权合同登记 图字：01-2013-5543 号。

图书在版编目（CIP）数据

固体氧化物燃料电池的动态建模与预测控制/（加）黄彪，（加）齐宇同，（加）莫希德（Murshed，A. M.）著；孙玉绣等译. —北京：机械工业出版社，2014. 11

（国际电气工程先进技术译丛）

书名原文：Dynamic modelling and predictive control in solid oxide fuel cells

ISBN 978-7-111-48484-4

Ⅰ. ①固… Ⅱ. ①黄…②齐…③莫…④孙… Ⅲ. ①固体－氧化物－燃料电池－系统建模②固体－氧化物－燃料电池－预测控制 Ⅳ. ①TM911. 4

中国版本图书馆 CIP 数据核字（2014）第 260967 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：顾 谦 责任编辑：顾 谦

版式设计：霍永明 责任校对：陈立辉

封面设计：马精明 责任印制：乔 宇

北京机工印刷厂印刷（三河市南杨庄国丰装订厂装订）

2015 年 1 月第 1 版第 1 次印刷

169mm × 239mm · 17. 5 印张 · 347 千字

0 001–2 500 册

标准书号：ISBN 978-7-111-48484-4

定价：79.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务	网络服务
服务咨询热线：(010)88361066	机工官网：www. cmpbook. com
读者购书热线：(010)68326294	机工官博：weibo. com/cmp1952
(010)88379203	教育服务网：www. cmpedu. com
封面无防伪标均为盗版	金书网：www. golden-book. com

本书特色

- 板式和管式SOFC的讨论以及详细、简化的SOFC动态模型。
- 系统地从单元层次到系统层次描述了单一模型和分布式模型。
- 为所有模型提供容易参考的参数以及复制结果。
- 所有理论都是通过生动的燃料电池应用实例演示,如最先进的无迹卡尔曼滤波、模型预测控制和SOFC系统辨识技术。

为中华崛起传播智慧
地址:北京市百万庄大街22号
邮政编码:100037

电话服务
服务咨询热线:010-88361066
读者购书热线:010-68326294
010-88379203

网络服务
机工官网:www.cmpbook.com
机工官博:weibo.com/cmp1952
金书网:www.golden-book.com
教育服务网:www.cmpedu.com
封面无防伪标均为盗版

策划编辑◎顾 谦

国际视野　科技前沿

国际电气工程先进技术译丛

《固体氧化物燃料电池的动态建模与预测控制》
《高级测量仪器与计算机I/O设计：精度界定和控制及其工业应用》（原书第2版）
《氢能源和车辆系统》
《纳米技术与能源》
《电能效率：技术与应用》
《自主移动机器人行为建模与控制》
《环境能源发电：太阳能、风能和海洋能》
《智能电网——融合可再生、分布式及高效能源》
《电动汽车融入现代电网》
《太阳能照明》
《超级电容器：材料、系统及应用》
《太阳能发电系统控制技术》
《ESD物理与器件》
《锂二次电池原理与应用》
《风力发电并网运行的无功管理》
《超大规模集成电路物理设计：从图分割到时序收敛》
《海上风电成本建模：安装与拆除》
《智能电网——设计与分析基础》
《功率理论与电能质量治理》
《混合动力电动汽车原理及应用前景》
《智能电网中的传导电磁干扰》
《风能转换技术进展》
《电机传动系统控制》
《储氢材料：储存性能表征》
《风力机技术》
《小型风力机：分析、设计与应用》
《通信系统高电压保护》
《磁性测量手册》
《ESD设计与综合》
《太阳能利用技术及工程应用》
《超级电容器的应用》
《小型风力机：分析、设计与应用》
《太阳电池、LED和二极管的原理：PN结的作用》
《风力发电系统 —— 技术与趋势》
《可持续电力系统的建模与控制：面向更为智能和绿色的电网》
《电力系统高级预测技术和发电优化调度》
《大规模储能技术》
《风电系统电能质量和稳定性对策》
《环境能源发电：太阳能、风能和海洋能》
《传热学：电力电子器件热管理》
《现代电力电子学与交流传动》
《功率半导体器件：原理、特性和可靠性》
《风能系统——实现安全可靠运行的优化设计与建设》
《储能技术》
《光伏系统工程》（原书第3版）
《光伏与风力发电系统并网变换器》
《先进电气驱动的分析、建摸与控制》

WILEY

Copies of this book sold without a Wiley Sticker on the cover are unauthorized and illegal

机械工业出版社微信服务号

上架指导　工业技术／电工技术／燃料电池

ISBN 978-7-111-48484-4

9 787111 484844 >

ISBN 978-7-111-48484-4　定价：79.00元